


```

Db 1801 ctggagataattttatcttcttgaatcttctgtgtcttcacatcttctacagtaat 1860
Oy 1801 CTGGAGATAATTTTATCTTCTTGAATCTTCTGTCTTCACATTTTCTACAGTGAT 1860
Db 1861 ataacaatagtaaggccgtaaataaagtcgaattagaagatccagctcttga 1920
Oy 1861 ATAAATCAATAGTAAAGGCCCTAAATAAAGTGATTTGAAAGATCCAGTCTTGA 1920
Db 1921 aaacacgtcttctgttaatgaagcagaattagtgtaatataaagtggaatcatt 1980
Oy 1921 AAACACTGTTCTGTGTAATGAAGCAGAAATTAAGTTGTAATATTAAGGTGATGCATT 1980
Db 1981 taagggaattaatcttcttctgtcgaagaagagatcattgattctgtcagtcaga 2040
Oy 1981 TAAGGAGATTACATCTTTATCTGTGTAAGAAGAGATCATTTGATTCGTACAGTCAGA 2040
Db 2041 acagacttggtgtgtcgaacagcttctgtgagaagaagctagtgtaataagttactga 2100
Oy 2041 ACAGACTTGGTGTGTGCACACAGCTTCTGTGAGAAAAGCTAGTGTAATAGTTAACTGA 2100
Db 2101 aagtttaactatlaaagaactaaatgcacattttagtgcatacttataaaagta 2160
Oy 2101 AAGTTTAAGTAAAGACTAAATGCACATTTATGATCTGATATTTAAAAAGTA 2160
Db 2161 atgtgaagcttcccttcttgaagtaatttctacagagtggtgaattgtgctt 2220
Oy 2161 AAGTGAAGCTTCCCTCTTGTATGAGTAAATTTATTCAGAGTTGTTGATTTGTCCTT 2220
Db 2221 ttaataaagtggaagctgcttctttaaataaataaagcggaattc 2269
Oy 2221 TTAATAAAGTGAAGCTTCTTTRTAAAAAAGCGGAATTC 2269

```

```

RESULT 2
LOCUS HSPROKINX 3018 bp RNA PRI 01-JUN-1995
DEFINITION H.sapiens mRNA for Ndr protein kinase.
ACCESSION Z53102
NID 9854169
KEYWORDS protein kinase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Carnivora; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3018)
AUTHORS Millward,T.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-1994) to the EMBL/GenBank/DBJ databases. Thomas
Millward, Friedrich Miescher-Institut, Basel, Postfach, 2543,
Switzerland
2 (bases 1 to 3018)
REFERENCE Millward,T., Cron,P. and Hemmings,B.A.
AUTHORS Molecular cloning and characterization of a conserved nuclear
TITLE serine/threonine protein kinase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 92 (11), 5022-5026 (1995)
MEDLINE 95381588
FEATURES
Source Location/Qualifiers
1..3018
/organism="Homo sapiens"
/clone="Human Ndr protein kinase"
/clone_11b="Human fetal brain/Stratagene 936206"
596..1993
/codon_start=1
/product="Ndr protein kinase"
/db_xref="PID:9854170"
/translation="MAATGTPCSMSNHTKERTVMTKVTLENYSLIAEEREMR
ORLEKVEEGLDDEERLRRAARETEFLRKRTLEDEESLKVGRAGFE
VRLOKDTGTVYAMKILRKADMLEVOVGHIAERDILVADSILAVYKMYSDOKL
NLVIMETPGDMTILMKKADLTLEETOFYIAETVLAIDSIHQIFIHDIKPDNL
LDSKGHVKLSDFLCTGLKKAHRIEFLRNINHLSPDFTONNSKRAKETWRNR

```

```

OIASTVGTDPDIAPFVFMOTGYNKLCDMMSLIVYEMILIGYPPCESTPOETRYKY
MMKKEITLFFPEPDISKAKDLIRFCECEBHRIGANGVSEISNSFSEVDENHIE
RPAISISEISIDFTSNFDEFPESDILKPTVAATSNHPTDYKKNQWFIWYTKRFEG
LTARGAIPSYMKAK"
BASE COUNT 908 a 640 c 672 g 798 t
ORIGIN
Query Match 1.5%; Score 35; DB 55; Length 3018;
Best Local Similarity 74.0%; Pred. No. 2.62e-05;
Matches 54; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Db 449 tttctcgtctgtgtcatatatactgctgtctccagaagaactgcacccaattac 508
CP 1419 TTTTAACTCTTTATCATTTTACCAATTCCTCTCAGAACTACTGTACCAATTAC 1360
Db 509 attcctaccaata 521
CP 1359 AGTCCATCACAATA 1347

```

```

RESULT 3
LOCUS SPAC16C9 14071 bp DNA PLN 22-OCT-1995
DEFINITION S.pombe chromosome I cosmid c16C9.
ACCESSION Z543365
NID g1019812
KEYWORDS MAX32 protein homologue; NAM7 protein homologue; NOT4 homologue;
serine/threonine protein kinase; zinc finger protein.
SOURCE fission yeast.
ORGANISM Schizosaccharomyces pombe
Eukaryotes; mitochondrial eukaryotes; Eumycota; Ascomycotina;
Archaeascomycetes; Schizosaccharomycetales;
Schizosaccharomycetaceae; Schizosaccharomycetes.
REFERENCE 1 (bases 1 to 14071)
AUTHORS Badcock,K. and Churcher,C.M.
JOURNAL Unpublished
2 (bases 1 to 14071)
REFERENCE Barrett,B.G., Rajandream,M.A. and Walsh,S.V.
AUTHORS Direct Submission
TITLE Submitted (12-OCT-1995) to the EMBL/GenBank/DBJ databases.
JOURNAL Schizosaccharomycetes pombe chromosome I sequencing project, Sanger
Centre, Hinxton Hall, Hinxton, Cambridge CB10 1RQ E-mail:
barrellesanger.ac.uk

```

```

COMMENT
Notes:
Details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, http://www.sanger.ac.uk/
yeastpub/sv/home.html)
protein coding regions (CDS) have been predicted with the help of
computer analysis using the Genetinder program in Pombase (an ACEDB
database) with additional predictions for the branch-acceptor sites
supplied by the program Sp3splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.
CDS are numbered using the following system eg SPAC510.01c. SP (S.
pombe), A (Chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
(complementary strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
Cosmid c16C9 is overlapped by cosmid c4G8 at the 5' end and by
cosmid c2G11 at the 3' end.
Location/Qualifiers
1..14071
/organism="Schizosaccharomyces pombe"
/strain="972h-1"
/chromosome="I"
/clone="cosmid c16C9"
/map="1L"

```

```

FEATURES
Source Location/Qualifiers
1..14071
/organism="Schizosaccharomyces pombe"
/strain="972h-1"
/chromosome="I"
/clone="cosmid c16C9"
/map="1L"

```

misc_feature 1..267
/note="overlap with 468"
complement(96..101)
/note="splice donor sequence, gtaagt"
complement(join<100..179,217..256))
/gene="SPAC16c9.01c"
/note="SPAC16c9.01c, partial spliced CDS, len: >39, similar to SW:MK32.YEAST P23060 MAK32 protein (46.7% identity in 30 aa overlap); pid:e204615"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019813"
/translation="MDPSPDVEFVTGWIIDELEDPDGCKIKIDIGAGSFG"
misc_feature
complement(186..193)
/gene="SPAC16c9.01c"
/note="splice branch and acceptor sequence, ctacactgaaag"
complement(211..216)
/gene="SPAC16c9.01c"
/note="splice donor sequence, gtaaaa"
complement(677..682)
/note="splice branch and acceptor sequence, ctacactgaa"
complement(join<777..1037,1137..1697,1747..1848))
/gene="SPAC16c9.02c"
/note="SPAC16c9.02c, unknown, len: 307, similar to SW:Y62.RH0RU P23139 hypothetical 25.8 kd protein (46.9% identity in 128 aa overlap); pid:e204616"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019814"
/translation="MNKQAEPIILGVIGSGFYDIPGDFVDSVNPITPWTGPPSPISG
IARTSGSLIPLARIGVGHGYITPREVSRNATLMAKLSGLVAIVSESAVGSILREDIIP
EDVILVQIIDITICARIPATPFEGCVAHVISEDPPDDDLIELISGCSNLKSGSLIP
HTRKRGDILVIGMESPASTSAESNLIRGSGASTLIMSVTIPAKIARAEALIAQVWC
MATYDQWRAEPEPVYEVETVMAHISNNKDNKITLLEAVKLEAPILLOGFLNRRESS
VONQIQNHKRRPDIKRLQGLFNPNTIPH"
misc_feature
complement(1038..1051)
/gene="SPAC16c9.02c"
/note="splice branch and acceptor sequence, ttaacaaaaaatg"
complement(1131..1136)
/gene="SPAC16c9.02c"
/note="splice donor sequence, gtaagt"
complement(1698..1713)
/gene="SPAC16c9.02c"
/note="splice branch and acceptor sequence, ctacactgatacag"
complement(1741..1746)
/gene="SPAC16c9.02c"
/note="splice donor sequence, gtagta"
3172..4668
/gene="SPAC16c9.03"
/note="SPAC16c9.03, unknown, len: 498, similar to SW:IKH0.YEAST P38616 hypothetical 59.1 kd protein (54.9% identity in 494 aa overlap); pid:e204617"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019815"
/translation="KMDNCGMGAVFQFSEPIASTILCEGCVPPNMAAMQDIT
MTMIDISGIPRESTYVNMHRCREGEVMOBPNNMMATLSEPELMATCIKINGLMQVLY
DAFNIETPESHRITVLYLQKEAFNTILQOSFOVEYVNNTOCPDCAITPIHAK
AOCVQKQALHRTFLYLEIILIKHRAHNTVNIKTRDGLIDITFGQAHAIKAVFL
SAVPIRVRGSGEILSEDPKSNNTAKYETISLEITFKDDVLCPTVAKAGNATL
LVACTCGGTGTFEIDPLTQICDMLPGLITIRTFPLADIDPELTETVAVDILGDTN
GKALADVLTIKSSDSGSHLTRHILGILAGNATVLAHLVAFNNNEVYDILGDSI
PVIYLYKTVYPTATKKNRMRKLTKIGQMAEDKVRKQREORYELFLQNLBEDPEL
KQGVNLYKRVAVNADVTDMDEVDIEDIQISVDILDDVEAMHI"
complement(join<5081..6317,6406..6585,6636..6688))
/gene="SPAC16c9.04c"
/note="SPAC16c9.04c, regulatory protein, len: 489, similar to SW:NOT4.YEAST P34909 general negative regulator of transcription subunit 4 (NOT4) (38.8% identity in 417 aa overlap); pid:e204618"

[illegible]

Matches	0:	Conservative	0:	Mismatches	55:	Indels	0:	Gaps	0:
Db	1136	ctgtagccatgaacaacattggttaggcaatctcagcttcgagctatgttgctcg	1195						
Cp	779	CTGTGGCGCATGGCGATACCTACTCGTACCAATTCACACCTCTCTTACCAAGAACACACTGTG	720						
Db	1196	gaatgactgcacatgttatatactcigcccccagcttcgtatataatgcttcgtgac	1255						
Cp	719	GAATGTGGTCAATGTTGATATACATCCGCCCCCGAGGTGGGAACATGAAGCTTTCTGCCC	660						
Db	1256	gcgactacaatgcagagacctccat	1280						
Cp	659	GGGAGCTTAACACGAGTCCCTCGAT	635						

RESULT	LOCUS	DEFINITION	ACCESSION	NID	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
4	CELBO228	38946 bp DNA	INV	30-MAR-1995					
	Caenorhabditis elegans	cosmid B0228.							
	U23168								
	976361								
	Caenorhabditis elegans strain-Bristol N2.								
	Caenorhabditis elegans								
	Eukaryotes: mitochondrial eukaryotes; eukaryote crown group; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Pseudocoelomata; Nematoda; Secernentea; Rhabdilita; Rhabditida; Rhabditina; Rhabditodea; Rhabditidae; Caenorhabditis.								
	1 (bases 1 to 38946)								
	Wilson, R., Almscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N., Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Showkneen, R., Smaildon, N., Smith, A., Sonhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstein, L., Wilkison-Sproat, J. and Woldman, P.								
	2.2 Mb Of contiguous nucleotide sequence from chromosome III of C. elegans								
	Nature 368 (6466), 32-38 (1994)								
JOURNAL MEDLINE REFERENCE	94150718								
AUTHORS	2 (bases 1 to 38946)								
TITLE	Leimbach, D.								
JOURNAL REFERENCE	3 (bases 1 to 38946)								
AUTHORS	Waterston, R.								
TITLE	Submitted								
JOURNAL COMMENT	Submitted (21-MAR-1995) Robert Waterston								
	Submitted by:								
	Genome Sequencing Center								
	Department of Genetics, Washington University,								
	St. Louis, MO 63110, USA, and								
	Sanger Centre, Hinxton Hall								
	Cambridge CB10 1RQ, England								
	e-mail: rtw@nematoe.wustl.edu and jess@sanger.ac.uk								

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is B0495, 8001 bp overlap; 3' cosmid is C06A8, 750 bp overlap. Actual start of this cosmid is at base position 4495 of CEBO228; actual end is at 6105 of CELC06A8

NOTES:

Coding sequences below are predicted from computer analysis, using the program GeneFinder (P. Green and L. Hillier, ms in preparation).

```

FEATURES
source
location/Qualifiers
1..3896
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
join(56..502,1135..1583,1735..2017,2219..2342,2390..2842,
2988..3076,3122..3241,3287..3355,3404..3472,3519..5054,
5181..5318)
/gene="B0228.4"
/map="111"
/note="coded for by C. elegans cDNA yk26d12.5; coded for
by C. elegans cDNA yk21b9.5; coded for by C. elegans cDNA
yk16f1.5; coded for by C. elegans cDNA yk32h8.5; coded for
by C. elegans cDNA yk16f1.3; coded for by C. elegans cDNA
yk32h8.3; coded for by C. elegans cDNA yk26d12.3; coded
for by C. elegans cDNA yk21b9.3"
/codon_start=1
/db_xref="PID:g726370"
/translation="MNDYERIDQLTGSGGASQNNOKISNPSGEAFNCTSNMYTDEPDR
DTCSSSLNKLNSNRHNSPMNHLSTYGCGVADALDLSLTTP.PQYVIRFGSITIDTPDR
SRNDKLTTFEYSSCKNLTLTTIIPDLVLVAAGVSTQALVLVAGMVCVCCCKC
CLCCLANGRGVRGQKQIOEPIIADDDGDLPELVGSHIPPYNKLKLSRSSH
VPYSPVSLVSSASADYKRLSMETSLVEDVLEPLRPLVYGKLDYIDKEGNPSC
CREGTRKPGFRPLRGSGSPDPDPSPSTLVAATYSAEVKIKDHKKLKEKKKK
MDKROSSGGGFEFRMGSGGSNNQONLADVDAESEKTKAPRODLQESERS
GGRHSHFEYRHQOPLMANTNDEDDYATIDARVSRNRRNSPMSPVNHVYDT
SGPLSRMTNESAFEDGRHAIQYRPAQGAAGPTTGAATSSVQDLDAITDLRLSTE
PSPSPRRALPKSLSSVQKQPIKRDAGGOLGSGVYTDONSVDTTIDDERAKDF
LRCGRSRLSPQSEKRNEROIQIROQSGPTNREETIEYERKOGPPVVRTYEGKL
KMEYIGADLITVDSKISAMVADTYNTKIKSTIGKSLILEEIKDGGSKYKLTIL
ENGTEKMERASDVPDPVFNKKDYLAVSCKLADPDESSESALHIEVEYEDVT
NLIKTYTIGERADYVLAEOURLHEDTAKTSPDPLEVEKTYDELQDKIYELD
PERKADHILKIDGRHFEEGALKRIRRETESTSEKPTVIMEPCAHAFADCVAKRD
DTSNYTYKIAVPLVHTITFLFKRSKMRQOAKAGYEMOEGHILRGETTKRKHKL
EEEEKOHAVAIQHVVEEVKATMKTQREVAEGGOYEMOEGHILRGETTKRKHKL
DSESEKREMEREAEGGOYAMOMEGERLGEKFRSGRGRHESSESNMAGNGSPT
LYDYKRESSSIEPATPEPTANNHSPYAPVPRPKIKENTTICCTISNOKASNAFLT
TKRYNTEKSGKTRFELGEBOAMMLCGFENOKSKEBYTSSROOKNETIKAFAGSKE
TENTITSTTIFHDADSFVAVSGSSSASNTATYGFGEKSEEMSNMYLYKATSESSSN
LSGETIKMKHIIORSSSEARFSEARFSEVQASAVMKNGVFERGSGTSTVAAEAAG"
join(6960..14594,15067..18774,19220..21034)
/gene="B0228.3"
/map="111"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726364"
/translation="MEESMAVECKRRSSSIDRASLKLASOEVCTGFNMTSGEST
RRTLIOTKTSVESWCLTKTSASTSDMPEDCORATYSPITREIASHREIYAAAG
ISTOSLERLKFVEIDIMENITMNVSSQKEDLSANQALTSAPNHLIDPSLGRVAPA
EODYSOLIMKRESESVFAPOLRASADEITREVALDSMSQLEDAAMSLTISVSKC
DLRTAPSNITIANTEIFYDYVAEEMKSVGWKRSRKHRSKFTSREBETIOGFMG
EDDEREVKTLKDRMVFSEFKSLSYEASSTSEVNSLKRADODKSEPTISOXSPKTI
VSEYAVGESSEKLOFQVMEKMDMSNTELEKEETASANTIRLAPQKCVGLGLGL
KAPQODSEVGTQOETRRADMAVSASLITLTSSNKSNTISDPERKAFNSLISV
ITSNBNSTIGTSESTTASISYODIPMLAASKIMISRNSEKLEKEIREYIQVEEF
RENTVDSDEKVALINERKIDVSYSLNLTALSMENETISOELVRTEAGVAVNIPSP
RELIVSSFKITSDLOQFPNVLKEMDSQSLPEAOEHALISKVNRVAPASELSITLP
GKLAPAMEOETETDLKIROQAARVYLVNVESSNKTISSESTFVRIENKATLTINY
GIITDLOSTLIGASENQTIMEDYTEKEELMSKTYVNDLSVLRKIEESGGEVVO
GFMTWASDKREKVGITCEKLSIHTHLLKATVIRPESLSTDIQDOBKSTHFWLTL
STRVAYOAPRISSESVYDHLITLYEKMDKIDIVEFYVNTISSVQALATSSOSSG
ILGLDLPAPPEETTSKEIREINTAKILVNLISLNTSTISDALNDIISSEAIFSN
ITIGMASNNLTISQOSTSSSGFNNVVELLEARVAVLOESKONQIOKVRSSSESLITLP
IMWASNEETKVALIYKEKLETVQAMKTLAOMTQVSNELIGSENLASIKITLIP
TRBYSIAFGINELVIOKMLVILKGEWSTILKGEWSTILKGEWSTISQVRLAPLASFNCDSI
LGRINAPQPOTEHSHLIEKMTAAANAAVKAAYEISSESSALKIPADEKATITML
AGLYASDVSSLCISSETFEGDRIPIRNOANISLGAVRSETLQAKREPIENOVQGF
RSTVANOERKASFIKAKOLDTMYDTMKVIAIOMISQTIQDGFSSVSSISFKNIDKRAA
REYVAAEFGIANESVOKALEVNLANKVEMSDISLPEKLAQNTISSVRLADSYNDSITII
GKLNAPPOSAYVDOQISEHQTLEVIINANKSAVSYSNDVSVLSRSEDEKAVYAGTA
NLISCLPASMSSTSNSEFERODILEPONANVLILGTASOFTFRNLQEPLOSIEVOGFW
STNSOKASISIVROKLTNVKYLDAKMKWATEIAPVQSLNNTISSEPTNEOSDPKSTR
ESTIASVGSSEYSEHOKTLEILSKSELSTIYUVAPEIATYQNTIKALSEPNCESITG
KLNAPDQASAMDIINNEQRKLTVANSIQSAIEVSIINSSLEKLEIVDSVLKISE

```

IYASHDITLASSCSOPFNILDOJNSTNOEKALFVETKPSMOTLTERLREPIREKOVQFPA
 TASSEJOEKFLEKEKETIHALOFKSAFSLSEVRODMATVOTLREAPLSTOQFPR
 ILAAGIGIEMEOLEOFJEDREDBMSHDIPCALRNDLNUYRINYADNAFNGNLS
 PRBPNDETVEYENKROJLOFLNLSRDIDJOSNOSLSTRSDRELYANSLTIS
 ENGLIDISAOVLOANSDDEKSFETPLNLLINPNVSETPESIOSMKTISQAE
 ASSTISERKLSMLOSEFVSAKOVSTLITDYKRIKPNVSEITFEDDAROVAAVS
 VQSEJLNETLESOLOEDMSOIKLSKROKSTLANSKLSANLUDITDQVRAVS
 OSEETLSEBSSVSEVNETKTOIIFPDYKREKSDERSKSTINLIGLTVNLAALQ
 VASFEJEDPOSIDJSTIPYKTOEDHROQTEOVQFENSKREHNEVYANRIVE
 LKSLINCVASCHSTIERSLSEKQODEFCATRGTEVQVNAFVQDSYKSLV
 GSNONEPEJAPKETIKSNKLILNLPYDITLAKTKEHVAKVEEOLQFOMFNK
 SANSTIEBMSQVSCIEEIEEYLPVNLVLAASANVETIIEKQSFHVLNMLSEI
 OHEBSRSALYRDIENLRLITDTPALIKQASSEVRYEIIIEKQSFHVLNMLSEI
 IKTOILMELYKVMONSTIAPEDLOAFONSODSOTVORVQVNAHEKIKLEETI
 RNLRLIMTAIPREVESTIAPEDLOAFONSODSOTVORVQVNAHEKIKLEETI
 KLINCEKCEVPTGCGYEMKPNDAITFTSMALGTLISCOLVYMASNLOSVDK
 YKRRAPKAEETVSEVNDHAFSLTALQDAEYVNSISTVSNQAAITVSDKLS
 KSTENMKRANVSANVANDHAFSLTALQDAEYVNSISTVSNQAAITVSDKLS
 VASEEDIAOVASPVNDVOTSMQSOOKLLENAKGTVEJREPEDEDTTAROR
 VYSSVOVAPAEESITQTEFLTRSELSEKLENAKGTVEJREPEDEDTTAROR
 EKVAVASAMTIVASEVLOKESIESFVSYESDNLVQCSMTTAPALVKSITQK
 VEVEYVAMTIVASEVLOKESIESFVSYESDNLVQCSMTTAPALVKSITQK
 TLEKRDQIEMSKITPORLEVEDESKAEQNDVETIIOGWGSLPAPIDQEEETULM
 KRASAPARILKAKETIELEDESEKAEQNDVETIIOGWGSLPAPIDQEEETULM
 SMKEYSKVESEVETLSTRSDRNSOQOKESPSLEMAITGKNGILKNAASRINKS
 PPRIDHAKAKAARONSTENMTGSKOXSPLKMGITVQVNAFNGNLS
 TLASQOVCYTORIDVSNITVASEPTVQSNKNETIGYQVNAFNGNLS
 IENLSQTSIKMNAJSDVTTVEKTLQHGNGTGDLSQVNAFNGNLS
 ISVRODREPEVAPKEDLNKSTINSTONLILSKREEESGITRSHEKTOQTL
 RHRSRSRSASRTVITAPNOEONDFKQVDSVAGSLSTIGVIREASQVQWAEI
 TSEILKLASNEVQVAVSETTNEFRKQVDSVAGSLSTIGVIREASQVQWAEI
 TOKLRLTYSRSKASMEASQVTOQKRDSDLASZYVARDLILKSSSVHATE
 QUSHELWAKRSSEHIEKQELERKESFSESQETISNOCHVIMVINDONDAAL
 MASTETBOKSDAONVETLAGTTLSTVRLGQDEKRIEHEVYIVASVILSTR
 ADHMSRSDAONVETLAGTTLSTVRLGQDEKRIEHEVYIVASVILSTR
 GENSDETULKSQFIDTALIEETTDMSDLILPCQDSVILSHKNOQVYDKAT
 ENVADQVTLLEKEVEQSDADVKKFEVEIEEVLIMIRRAELISMAKASVIO
 HENKRSERKAEQOXLKALNLEIKINVALEHEVYIVASVILSTR
 DKSHERKEVQESONMNLSTNAEVELINDLEDSVITASOVSADMSAFVASATV
 LINDSITAKAPPEIGIKESQSNVYKVALKVEKSTENHVAVASFROIEETMG
 DEISANLEAITPTSGVAKIKSTARSMSFETSHASEETATVADVAROAYS
 TAARKAAPETPTSOIKLEVGAKIGISLIMGEITDITDATOVLVAKSIOIS
 SESEVARNOTLEMPKEKGVILINVONKECEENPASTISIDYHKSKEITPTI
 VSEKLETKERVLRFTVYEGSVGRVNAYNLHRSKSLPTIDYVREYVRIKALP
 LFTISAGDESVYVSTHOLEKTFVREDEITDITARGDENKIKKEDEYVREKARVE
 LIGNLNPEJICDITSMKTPRGKVLVLEPEPDEJCLFMSNOLQDRE
 16445... 25656 26650 26780 26826... 269000 26948... 27083,
 27129... 27228 27467... 27586)
 27129... 27228 27467... 27586)

/note--coded for *C. elegans* CDNA yk139c1.5; coded for by *C. elegans* CDNA yk45a10.5; coded for by *C. elegans* CDNA yk45a10.3; coded for by *C. elegans* CDNA yk21e5.3; coded for by *C. elegans* CDNA yk35c1.3; coded for by *C. elegans* CDNA yk45a10.3

/db_xref="PID:9726363" /translation="MRKRESQIVGVCAVVPBENLSAKLDEKSGGCEYLTTR
AAGEYEMISSILSNAPQVAKYKQKGLQAVAKDELPLETTEBSVANNFVEYLPENL
FKIATCTANNNAPPMFMRHLECDITVAFNMNMKQDFSLKLTISNGISKTL
SCYAELEHTNPERRPHESDITQVADNLEKGLNLFELNMKLSLMBLSR
ODDSRIEHTVKNDRNANIKFPEYTESSEKQTSASIEDITEKEIRITIDYLRQOG
OKFISTDAEENETLLHARIDSKITFYHYTHLILINSASAHQIISTGASLEHNTS
TQLSRPAEMNETELLVDKNVQVPMYARLKDCEVENLHPIYRPDDIFELDITISV
ARNGKREKRCRKSAGEKEFEDEISRSIKDDIIDITGVNGOEPSTFOTISV
TAVSDOMASLGERKAEIKALYQNSVNGINVEKMTAEVEYETISEQRKNDKEED
YASVDAAGSGLSTLNASQSSSVNNLIPRPSQISIKETFTQITOTIPIATSRK

	Query Match	Similarity	Score	Length
Best Local	65.0%	Pred. 61.8E-03	38946	
Matches	67	Conservative	0	Mismatches 36; Indels 0; Gaps 0;
Db	33432	catcaacttggttcacatgtgtactacgcgaactgacatgcagatccagtaaatat	33491	
Cp	251	CATCAGATGGCTTGGCAATAGTAGATTCACATATTTTCAGTCTCTCTCTTAAATTT	192	
Db	33492	ttagactctcaaacacactctcccaataatcccaactgtgc	33534	
Cp	191	CTGATCATCCAGGCCCTTTCACCACTTATTCACATTTTCATCC	149	

LOCUS	5	1117 bp	DNA	MM	02-MAR-1995
DEFINITION	MHC 67-mycosin heavy-chain 67 (exons 31-33, and 40) [cattle, Genomic, 1117 nt, segment 1 of 2].				
ACCESSION	S730781				
NID	S7307				
KEYWORDS	9685132				
SEGMENT	1 of 2				
SOURCE	cattle.				
ORGANISM	Bos sp.				
REFERENCE	1 (bases 1 to 1117)				
AUTHORS	Young,R.B., Hsieh,M.Y., Hudson,J.R,Jr., Richerf,H.E. and Scott,M.				
TITLE	Expression pattern and partial sequence analysis of a fetal bovine myosin heavy-chain gene				
JOURNAL	J. Anim. Sci. 72 (4), 903-910 (1994)				
MEDLINE	94284221				
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gdlbseq 155276] from the original journal article. This sequence comes from Fig. 3.				

FEATURES	SOURCE	location/Qualifiers
		1..1117
		/organism="Bos sp."
		/note="cattle"
		join(221..367,455..626,713..837)
		/partial
		/gene="MHC 67"
		/note="Only first segment of protein sequence is bound: mismatch(70 K->N) : Description: myosin heavy-chain 67, MHC 67: Method: conceptual translation supplied by author. This sequence comes from Fig. 3. Author-given protein sequence differs from the conceptual translation for reasons explained in citation. Author-given protein sequence is in conflict with the conceptual translation. /codon.start=1
		/product="myosin heavy-chain 67"
		/db_xref="pid:g685134"
		/translation="KLLAQRLODAEEHVEVNVNKAASLEKTKORLQNVDELMDKKQ RNFDELSMKRQYETQELDASQKSSSLSTELFKYNAAEESDLETLTKRNNNN KRLQGEIDLTQELQAEQGGKIHLEIKIKQVDEKCEIDALAEAE"
BASE COUNT	350 a	243 c 236 g 288 t
ORIGIN		

Query Match	1.3%	Score 29	DB 34	Length 1117
Best Local Similarity	70.4%	Pred. No.	8, 22e-02	
Matches	50	Conservative	0	Mismatches 21
		Indels	0	Gaps 0
Db	894	tgtgtgaaccttggaacatataaaccttgaatgaaggccaattgacagcgcatacaaaa	953	
QY	1350	TGATGGAGCTCAAAATTGTCACAGTATTTCGTGAGGCGCAATTTGTAATAATGCATCAAAA	1409	
Db	954	gtcctataaaaat	964	
QY	1410	GACCTTAAAAAT	1420	
RESULT	6			
LOCUS	HUMRETBLAS	180388	bp	DNA
DEFINITION	Human retinoblastoma susceptibility gene exons 1-27, complete cds			
ACCESSION	U11910			


```

Intron /number=14
       76487..76688
exon /number=14
       76689..76920
Intron /number=15
       76921..77000
exon /number=15
       77001..77077
Intron /number=16
       77078..78082
exon /number=16
       78083..78279
Intron /number=17
       78280..149997
       /number=17
variation 99426
       /standard_name="Restriction Fragment Length Polymorphism"
       /note="XbaI RFLP"
satellite 102921..102959
       /standard_name="microsatellite RB17"
       123912..123501
repeat_region /standard_name="Variable Number Tandem Repeats (VNTR)"
       /rfl_type=tandem
       149998..150116
exon /number=18
       150117..153207
Intron /number=18
       153208..153353
exon /number=19
       153354..156692
Intron /number=19
       156693..156838
exon /number=20
       156839..160729
Intron /number=20
       160730..160834
satellite 160835..161996
       /standard_name="microsatellite RB1.20"
       160730..160834
exon /number=21
       160835..161996
Intron /number=21
       161997..162110
exon /number=22
       162111..162203
Intron /number=22
       162204..162367
exon /number=23
       162368..170371
Intron /number=23
       170372..170402
exon /number=24

```

Note: remainder of annotations omitted.

```

Query Match 1.38; Score 29; DB 60; Length 180388;
Best Local Similarity 62.9%; Pred. No. 8,228-02;
Matches 83; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

Db 7263 ctgttaataataataatgaatcctgattattattatagaataaattcttataag 7322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 1509 CTTTATATATATATCTTTGTATGCATCTCTGTATGCTTTAAAGAAATGCTAGATC 1450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7323 ttgaattactaa-ataaagactatgaacattttatgaccttgatgatattgctaga 7381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 1449 TAGAGTTCCAGACCAAGACGACGTCGTAATTTTAACTCTTTGATGCAATTTACCAA 1390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7382 ttacactcaaga 7393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 1389 TTGCCTCCAGA 1378

```

```

RESULT 7 HS117E6F 127 bp DNA PRI 19-OCT-1995
LOCUS

```

```

DEFINITION H.sapiens Cpg DNA, clone 1176, forward read cpg1176.fcla .
ACCESSION Z58950
NID 91030863
KEYWORDS Cpg island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
1 (bases 1 to 127)
Macdonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
Direct Submission
Submitted (16-OCT-1995) to the EMBL/GenBank/DBJ databases. The
Sanger Centre, Hinxton, Cambridgeshire, CB10 1RO, England. E-mail
contact: humpd@sanger.ac.uk
2 (bases 1 to 127)
Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
Purification of Cpg islands using a methylated DNA binding column
Nature Genet. 6 (3), 236-244 (1994)
94282070
COMMENT Vector: pGEM-5zf(-).
FEATURES
    source
        1..127
            /organism="Homo sapiens"
            /dev_stage="adult"
            /sex="male"
            /tissue_type="blood"
            /clone_lib="CGI-1"
            /clone="1176"

```

```

BASE COUNT 52 a 28 c 14 g 31 t 2 others
ORIGIN

```

```

Query Match 1.28; Score 27; DB 50; Length 127;
Best Local Similarity 73.6%; Pred. No. 9,788-01;
Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db 73 antgaattggtacacatttctgagacacantttagcaaaaccccaaa 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 1357 ACTGTAAATTGGTACAGATTTCTGAGCGCAATTTGTAATAATGCATCAAA 1409

```

```

RESULT 8 HS117E6F 127 bp DNA PRI 19-OCT-1995
LOCUS
DEFINITION H.sapiens Cpg DNA, clone 1176, reverse read cpg1176.rcla .
ACCESSION Z58951
NID 91030864
KEYWORDS Cpg island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
1 (bases 1 to 127)
Macdonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
Direct Submission
Submitted (16-OCT-1995) to the EMBL/GenBank/DBJ databases. The
Sanger Centre, Hinxton, Cambridgeshire, CB10 1RO, England. E-mail
contact: humpd@sanger.ac.uk
2 (bases 1 to 127)
Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
Purification of Cpg islands using a methylated DNA binding column
Nature Genet. 6 (3), 236-244 (1994)
94282070
COMMENT Vector: pGEM-5zf(-).
FEATURES
    source
        1..127
            /organism="Homo sapiens"
            /dev_stage="adult"
            /sex="male"
            /tissue_type="blood"
            /clone_lib="CGI-1"
            /clone="1176"

```

```

BASE COUNT 30 a 14 c 29 g 53 t 1 others

```


[illegible]

SOURCE	Candida parapsilosis.
ORGANISM	Candida parapsilosis Eukaryote; mitochondrial eukaryotes; Eumycota; Ascomycotina; Hemiascomycetes; Saccharomycetales; Saccharomycetaceae; mitosporic Saccharomycetaceae; Candida.
REFERENCE	1 (bases 1 to 675)
AUTHORS	Nosek,J. and Fukuhara,H.
TITLE	Mitochondrial transfer RNA genes of the yeast Candida parapsilosis
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 675)
AUTHORS	Nosek,J.
TITLE	Direct Submission
JOURNAL	Submitted (02-NOV-1993) to the EMBL/GenBank/DBJ databases. J. Nosek, Institut Curie Biologie, Centre Universitaire, Batiment 110, 91405 Orsay, FRANCE
FEATURES	Location/Qualifiers
SOURCE	1..675
tRNA	/organism="Candida parapsilosis" /strain="SR23"
CDS	9..80 117..347 /gene="trnI" /product="transfer RNA-Ileu" /note="ATPase" /codon_start=1 /transl_table=12 /product="ATP synthetase subunit 9" /db_xref="PID:g473030" /translation="MQIALAKKIGASIRTLGCGAAGTALGVALLINSTRNSPLR STLEPQAIIIGFALSEACGLFCMLISFLLYAV" 485..555 /gene="trnN" /product="transfer RNA-Asn"
BASE COUNT	223 a 70 c 83 g 299 t
ORIGIN	
Query Match	1.2%; Score 27; DB 44; Length 675;
Best Local Similarity	74.5%; Pred. NO. 9.78e-04;
Matches	41; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Db	370 ttataatataattatttatttaagaatttcctatatatatatatatat 424 Cp 1542 TTATTATTATACATCATTTATTAAACCCTCTTTATTATATATCTTTGTAT 1488
RESULT	11
LOCUS	ARTUEAP 1773 bp RNA PLN 10-APR-1992
DEFINITION	A.thaliana mRNA for leucine aminopeptidase.
ACCESSION	X63444
NID	g16393
KEYWORDS	aminopeptidase; exopeptidase; leucine aminopeptidase. thale cress.
SOURCE	Arabidopsis thaliana
ORGANISM	Eukaryote; mitochondrial eukaryotes; Chlorophyta/Embryophyta group; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Magnoliopsids; Dilleniidae; Caprariales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 1773)
AUTHORS	Bartling,D.
TITLE	Direct Submission
JOURNAL	Submitted (06-DEC-1991) to the EMBL/GenBank/DBJ databases. D. Bartling, Lehrstuhl fuer Pflanzenphysiologie, Ruhr-Universitaet Bochum, Postfach 10 21 48, 4630 Bochum 1, FRG
REFERENCE	2 (bases 1 to 1773)
AUTHORS	Bartling,D. and Weller,E.W.
TITLE	Leucine aminopeptidase from Arabidopsis thaliana. Molecular evidence for a phylogenetically conserved enzyme of protein turnover in higher plants
JOURNAL	Eur. J. Biochem. 205 (1), 425-431 (1992)
FEATURES	92209533 Location/Qualifiers
SOURCE	1..1773 /organism="Arabidopsis thaliana"

	RESULT	15			BCT	15-FEB-1991
LOCUS	RRCYTBCI	5550 bp	DNA			
DEFINITION	R. rubrum cytochrome bcl-complex genes (petA, petB, petC).					
ACCESSION	X53387					
NID	g46382					
KEYWORDS	cytochrome b; cytochrome bc-1 complex; cytochrome c1; petA gene; petB gene; petC gene; Rieske Fe-S protein.					
SOURCE	Rhodospirillum rubrum. Rhodospirillum rubrum. Eubacteria; Proteobacteria; alpha subdivision; Rhodospirillum. 1 (bases 1 to 5500)					
REFERENCE	Majewski,C. Submitted (05-NOV-1990) to the EMBL/GenBank/DDBJ databases . C. Majewski, RUHR UNIVERSITAET BOCHUM, LEHRSTUHL F BIOCHEMIE DER PFLANZEN, RUHR UNIVERSITAET, POSTFACH 102148, 4630 BOCHUM 1 2 (bases 1 to 5550) Majewski,C. and Trebst,A. The pet genes of Rhodospirillum rubrum: cloning and sequencing of the genes for the cytochrome bcl-complex Mol. Gen. Genet. 224 (3), 373-382 (1990)					
AUTHORS	JOURNAL MEDLINE					
TITLE	FEATURES					
JOURNAL	source					
MEEDINE						
FEATURES						
Source						
1..5550	Location/Qualifiers					
/organism="Rhodospirillum rubrum"						
/strain="Fr1"						
/clone_lib="Partially digested Sau3A fragments of genomic						
Rs. rubrum DNA in EMBL3."						
/clone="EMBL-pet 11.1, EMBL-pet 4.1"						
<3..>770						
/gene="ORF1"						
3..770						
/gene="ORF1"						
/codon_start=1						
/transl_table=1						
/db_xref="PID:g46383"						
/translation="MLDAPGRMPDPAGEGCALRLLDGSKDIVLDALLRNAERG						
KGPYRAERTARMRALIETALGWATPFFPLGGRDALACGKPAVGEMPACADAGR						
AAAPGGGTICRAPHSGSPLGCGTPTRNPRSRAMNRYSRDRSPGVLRPACMAOGLAE						
HOGGDGDDVERAHGNNHRRNPDAQYGALVDLGSNNRALSQAQQHYVLKVSFGKKRARL						
RCODDGAFGVEPEDEIRPGMTHTKSAAFELVHGSA"						
prim_transcript <3..>770						
/gene="ORF1"						
1250..1253						
prim_transcript <1261..>1812						
/gene="petA"						
<1261..>1812						
/gene="petA"						
1261..1812						
/gene="petA"						
/codon_start=1						
/transl_table=1						
/product="Rieske fes protein"						
/db_xref="PID:g46384"						
/translation="MAEAENASTRGGESSRRDELITRTAVGAYVALWVPPIQDM						
NPADDTIALASTEYDVSAIAEGCAITYTTRKRKPFTFYARROKELEYVARAYDPSLRDP						
OTDDRARVOAQOMLVWGVCYTHLCIPJGOAKDGKDFDWFCPCBGSHSDASGRIRK						
GPAFLNLPPVPYAFTDDTYVLIG"						
1811..1814						
prim_transcript <1826..>3043						
/gene="petB"						
1826..3043						
/gene="petB"						
/codon_start=1						
/transl_table=1						
/product="cytochrome b"						
/db_xref="PID:g46385"						
/translation="MYTPTRMNNKALKMFDERLPYLVAHKELVVYPARPNTLYFMNF						
GSLAGIMIIIMATGIIFLAISYTAHVDAEDSVIRINDNYTGILNKYMANGANSFE						
LIVTVWHFRGLYGISIKPREVLMMDLVILLMMATAFNGYVLPWCQSFWGATVTT						
NLSFAIPVAGDDIYTLTMGFESVDNPLNRFSLATLEPHMLEAVVELLHMWLFHHVKKS						

[illegible]

Tue Feb 11 11:27:24 1997

US-08-674-311-1-01.rge

Page 12

OY 446 ATTTGTCAGTTCATTGACGACC 472
Search completed: Mon Feb 10 22:51:01 1997
Job time : 1616 secs.

QY 1053 ATGGGAAAGCAGCAGCTTTCATGCTGCTATCAAGATATTTGTAAGAAAGAC 1112
 Db 2450 aagacattggtglat-agaagctcccaatgatttagaacaattcaaaatacagaaga 2508
 QY 1113 AAGACATTTGTTGATTAAGAGACTCTGATGATTTAACAACCTTCAAAATACGAAAGAA 1172
 Db 2509 aagcaaatgagctagtaa-catgtgggaaaaaataatcatcatttaaggaggaaaaaa-- 2565
 QY 1173 AAGCAATGACTGTAACATGTAAGGAAAAAATATTACATTTTAAGGGGAAAAAATAAA 1232
 Db 2566 cccacacattctctcccccatttaattggcaacaataaagggtggagggaatctc 2625
 QY 1233 CCCACCATTTCTTTCTCCCTATTAAATTGCAACAATAAAGGTGGAGGTAATCTC 1292
 Db 2626 tacttctctactgcaagaagatgaggaagaatggagactcttggtattattatga 2685
 QY 1293 TACTTCTCTATCTGCAAGAAATGAGGAAATGGGACTCTTGGTTATTATTGA 1352
 Db 2686 tggcactgtaaatggtacagatattctggaggcaattggttaaatgcatcaaaagac 2745
 QY 1353 TGGGACTGTAAATGCTACAGTATTCTGAGGCAATTGTGTAATAATGCATCAAAAGAC 1412
 Db 2746 ttaaaatacggagctacmnmr 2768
 QY 1413 TTAATAATACGAGCTCCTTTGG 1435
 RESULT 2
 ID 115167 standard: DNA: 2763 BP.
 AC 115167:
 DE 23-JUN-1996 (first entry)
 KW Human; methyladenosine-phosphorylase gene.
 KW Human; methyladenosine-phosphorylase.
 KW Interferon-alpha; CDK41; tumour suppressor; chromosome-9p21;
 KW cyclin-dependent kinase-4-inhibitor; probe; mutation;
 KW leukemia; melanoma; diagnosis; dysplastic nevus syndrome; glioma;
 KW non-small cell lung carcinoma; cancer; gene therapy; antisense;
 KW ribozyme; antibody; imaging; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 254..421
 FT /tag- b 422..615
 FT /tag- c 616..720
 FT /tag- d 721..963
 FT /tag- e 964..1203
 FT exon
 FT /tag- f
 FN WO9528169-A1.
 PD 26-OCT-1995.
 PE 12-APR-1995; 004655.
 PR 14-APR-1994; US-227800.
 PA (REGC) UNIV CALIFORNIA.
 PI Carson DA, Nobori T.
 PI WPI: 95-313630/48
 DR Cyclin-dependent kinase inhibitor gene, related vectors and
 PT antibodies, useful for diagnosis, assessing predisposition and
 treatment of cancers
 PS Example 1; Page 96-101; 129pp; English.
 CC The sequence encodes a methyladenosine-phosphorylase, and is
 CC located at chromosome-9p21. A cyclin-dependent
 CC protein-kinase-4-inhibitor (CDK41) tumour suppressor gene
 CC (115157-58) is located between this gene and an interferon-alpha
 CC gene cluster, and has been isolated by chromosome walking. The
 CC CDK41 gene, probe and primer derivatives and the gene product may
 CC be used in diagnosis of cancer, particularly melanoma (especially
 CC dysplastic nevus syndrome), glioma, non-small cell lung carcinoma
 CC or leukaemia. The gene may also be used in cancer gene therapy, or
 CC in antitumour antisense oligonucleotide or ribozyme construction.
 CC Antibodies against CDK41 may be used in diagnosis or in vivo imaging.

only
 base
 NTP

SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T;
 Query Match 20.5%; Score 466; DB 19; Length 2763;
 Best Local Similarity 99.0%; Pred. No. 5,206-267;
 Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
 Db 2270 agaatggcccgatcttctggtttatattacaaagacataaagtagcatggtccag 2329
 QY 933 AGAATGCGCCGATTTCTGTTTATTAACCAAGACATTAAGTAGCATGCGCCAG 992
 Db 2330 agaaagaagacattcattccagtcatttgggaattccgtcttaacttgaaaaaa 2389
 QY 993 AGAAAGAGACATTTCTATTCAGTCAATTTGGGAATTCCTGTTAACTTGAAAAAT 1052
 Db 2390 atgggaagagcattgacgttttacccttccctataaagaagatggttgaagaagc 2449
 QY 1053 ATGGAAAGACATGACAGCTTTCAATGCCCTATCAAGAGATGTTTAAGAAAGAC 1112
 Db 2450 aagacattggtglat-agaagctcccaatgatttagaacaattcaaaatacagaaga 2508
 QY 1113 AAGCATTTGTTGATTAAGAGACTCTGATGATTTAGACAACTTCAAAATACGAAAGAA 1172
 Db 2509 aagcaaatgagctagtaa-catgtgggaaaaaataatcatcatttaaggaggaaaaaa-- 2565
 QY 1173 AAGCAATGACTGTAACATGTAAGGAAAAAATATTACATTTTAAGGGGAAAAAATAAA 1232
 Db 2566 cccacacattctctcccccatttaattggcaacaataaagggtggagggaatctc 2625
 QY 1233 CCCACCATTTCTTTCTCCCTATTAAATTGCAACAATAAAGGTGGAGGTAATCTC 1292
 Db 2626 tacttctctactgcaagaagatgaggaagaatggagactcttggtattattatga 2685
 QY 1293 TACTTCTCTATCTGCAAGAAATGAGGAAATGGGACTCTTGGTTATTATTGA 1352
 Db 2686 tggcactgtaaatggtacagatattctggaggcaattggttaaatgcatcaaaagac 2745
 QY 1353 TGGGACTGTAAATGCTACAGTATTCTGAGGCAATTGTGTAATAATGCATCAAAAGAC 1412
 Db 2746 ttaaaatacggagcgt 2761
 QY 1413 TTAATAATACGAGCT 1428
 RESULT 3
 ID 092813 standard: DNA: 2763 BP.
 AC 092813:
 DE 05-DEC-1995 (first entry)
 KW Human kinase.
 KW Human; methyladenosine-phosphatase; malignancy; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 254..421
 FT /tag- a 616..720
 FT /tag- b 964..1203
 FT /tag- c 1640..1762
 FT /tag- d 2272..2310
 FT exon
 FT /tag- e
 FN WO9516233-A.
 PD 06-JUL-1995.
 PE 22-DEC-1994; 014920.
 PR 29-DEC-1993; US-116855.
 PA (CIBA) CIBA GEIGY CORP.
 PA (REGC) UNIV CALIFORNIA.
 PI Carson DA, Nobori T, Takabayashi K.
 PI WPI: 95-246398/32.
 DR Detecting methyladenosine phosphatase in mammalian cells - by
 PT hybridisation with specific oligonucleotide for detecting malignancy,
 PT also new nucleic acid, expression vectors, derived polypeptide(s) and
 PT antibodies

D	b	597	aycngsnndnnndnnndnmrrysnnndvkgmanhmsnsghsksncvdsrn	656
Q	y	1012	ttccagctcatitttggaatctcgtctaactgaaaaaaatatggaaaagactgcagct	10711
D	b	657	vhkntdygnasrztamdnmnanyskkntnnnsgnnntctgmnaayngnnnnn	716
Q	y	1072	ttcagatcccttgcctatccaaagagttga-ttgaagaaagcaagacattgctgatat	1130
D	b	717	nanrsngnyngndnsknvnykvrngnrtnrnsndtlnnnnnvnmnrwandnau	776
Q	y	1131	gagacccctcgatgatttttagcaacttcaaaatfacgaagaaaaagcaaaagactataaa	1190
D	b	777	ndngnnkgnrnrnkngtsgndnnnmnyannnnkvnrtlnaynnkrtkanayn	836
Q	y	1191	cnrtgggaaaaaataattactttttaagggggaaaa-aaaaaccacactctctct	1249
D	b	837	nnnhsvannhkgntlvnanadsvtlynsdnygtlensanbstmmvyltnndytcda	896
Q	y	1250	ccccctattaaatttgcaacataaaggctgaggttaattctactttcttactgccc	1309
D	b	897	ndhndykvntngdymvvsngnrngnrhannnarmanannadvssnrnrhndntr	956
Q	y	1310	aaagaattgtgagagaaagaaatggcactcttggctttattttgaagcactgtaaatgt	1369
D	b	957	rngvthgncvgaavynkhmrncngdvtvtlaasrnsngnanhvsctlkdanngcn	1016
Q	y	1370	acagatttttcgagagcaattgtrtaaatgatcattcaaaagacttaaaaaatcgagagct	1429
D	b	1017	nnnrtdvmnkgykmyr	1035
Q	y	1430	ctttgctgctgggaaactct	1448

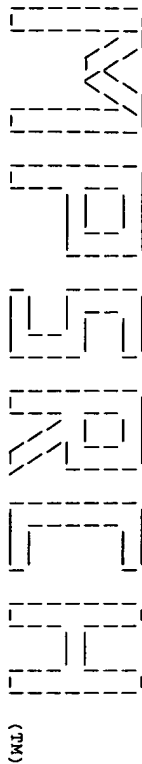
ID	010572	standard; DNA; 1047 BP.
AC	010572:	
DT	06-APR-1991	(first entry)
DE	Human Natriuretic Peptide Receptor B.	
KB	NPB; ANP; BNP; CNP; Kidney failure; heart failure; protein kinase.	
KW	hyperaldosteronism; glaucoma; guanyl cyclase.	
OS	Homo sapiens.	
FX	key	Location/Qualifiers
FT	peptide	1..22
FT	/label= signal sequence	
FT	protein	12
FT	/label= mature NPB	
FT	Domain	23..455
FT	/label= extracellular domain	
FT	/note= "binds natriuretic peptides A,B and C"	
FT	Domain	456..456
FT	/label= transmembrane domain	
FT	Domain	479..1047
FT	/label= cytoplasmic domain	
FT	/note= "GC and protein kinase activity"	
FT	Modified -site	24..26
FT	/label= N-glycos -site	
FT	Modified -site	35..37
FT	/label= N-glycos -site	
FT	Modified -site	161..163
FT	/label= N-glycos -site	
FT	Modified -site	195..197
FT	/label= N-glycos -site	
FT	Modified -site	244..246
FT	/label= N-glycos -site	
FT	Modified -site	277..279
FT	/label= N-glycos -site	
FT	Modified -site	349..351
FT	/label= N-glycos -site	
FT	Modified -site	600..602
FT	/label= N-glycos -site	
PN	W09100292-A.	
PD	10-JAN-1991.	

PF 22-JUN-1990; U03586.
 PR 23-JUN-1989; US-370673.
 PA (GETH) GENENTECH INC.
 PI Chang M, Goeddel D, Lowe D:
 DR WPI: 91-036711/05.
 DR N-PDB: Q10324.
 PT Naturalistic protein receptor B - for diagnosis and treatment of
 PT kidney failure, heart failure, hyperalosteronism, glaucoma etc.
 PS Claim 3: Fig 1; 49pp; English.
 CC The sequence was derived from the DNA encoding naturalistic peptid
 CC receptor B, NRRB, having guanylyl cyclase (GC) activity and proteinase
 CC kinase activity. The DNA can be inserted into expression vectors
 CC for the prodn. of the protein, opt. after being mutated to productu
 CC NRRB analogues. The protein has a mol wt. of 115 kD (calculated M
 CC 114.952). The protein (or variants) can be used in treatment of
 CC naturalistic peptide disorders, and also to isolate peptides using
 CC affinity chromatography. Antibodies with affinity for NRRB can
 CC also be prep'd.
 CC Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;

Query Match	3.3%	Score 76;	DB 2;	Length 1047;
Best Local Similarity	9.7%;	Pred. No. 1,15E-26;		
Matches	89;	Conservative 249;	Mismatches 567;	Indels 11; Gaps 11
D114	hvrnnnnhtagaaagnsakndhyrtnvrtgtsaakngnvnthghnnataraannnda	173		
CP931	CAGGTATGAGGGGTTTGTGACCATTC-TGTGACACCTATCTAG-GTATGGTAGAGAC	874		
D174	rtddrrhlytlngnnannngsnnsnhvyrnmngnnnathnrgsnvnyngnnmn	233		
CP873	AGTAAGCTTTGGCTTATTAAGCTTTCTTCAAGGGCTTTAAAGCCGGGTCAACCGAA	814		
D234	hnnnnnannrntngdyvnyndvn-gnsrtaqrtrtgyrwnrdtltmnaennant	292		
CP813	ACTGCTTCTCGTGGCTCCCTTCGAGAGCATATCTGTCCGATCGCATCTGGCTGA	754		
D293	vnyvlyrnnnnnnnnnnnnrnnrarnrdngvngnsnnnnag-cnydgnnnyannnt	351		
CP753	CAAAATTCACAGCCTCTTACGAAGAACCCTCTGGAACGTGTGCATGTGTAATACATCC	694		
D352	nnngtrtndgrnvnhkmngryhyvrgyavvdkndrtdnvnwangdndsgdnnaahy	411		
CP693	GCCCCCGAGGTGGGAAATGAAGCTTCTCCCGGAGCTAAAGAGGTCCCTCGATT	634		
D412	sganhnwvgyrnnnvkgansndnnncandndngsktnstbnvangeitnmmg	471		
CP633	GTGACCATGTCCCTTTGATGGTCACC-GGAGTCTTACGCTTCTTACAGTCTCTATAG	575		
D472	vssnnnnrkmmnnknnasnmwrrvnnnnnngnsrlykbgsgsrntnsngssygmtah	531		
CP574	AACTCTCTCGTTTGGGGCAAAAGGCTCAGCATTGTGAATA-TGGCACATCTCTCTG	516		
D532	gkyynnantghkngvnnvkhvnrkrrnnntnnnnnnkhmrdvnnhntctngacdnm	591		
CP515	CACAGAAATACTCCATCATGAAGAGACTGAGTGTCTATAGTGTCTCTCATTAATACT	456		
D592	ncnvtngyrnsndnnndsnnd-wmrysnndnv-kymanhhnsmnsghsksn	649		
CP455	GATCAATAATGACAATATCGCGGGCTGAATCTCTCCCTCAAGAGCCACAAGCTGTGG	396		
D650	cvdsrtrvnhktdygnasnrstannndnnnyakknkntannnsgnnntltgmadaivs	709		
CP395	TCACATAGACATGTGTACAGCCCTCTTCCTTCAAGGCCACAGATTTGCGCTGTAGT-TG	337		
D710	ngnnnnnnnacsrgnnyngndnsnknnvkvngnrnrnrnsndrtlnnnnnvnnmmxk	769		
CP336	ACCTTGAAAGCATGATGGTGGTCTCTTCATCTCCTTGCAGAGAGAGCATATACA	277		
D770	andannrndgnkngrnrnkngsgnsndnnnnnmn-yannnhkvnrrth-aynnnk	827		
CP276	TTTTTTATCTTCCCAAAATAAGCATAGATGGCTTGCCAAAATGAGATATCCACATAT	217		
D828	rkanannymnnhsvannnkrgntlvnanandsvtlynsdnygtntsansntmmvvlmn	887		

[illegible][illegible]

Search completed: Mon Feb 10 22:55:11 1997
Job time : 231 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MSPrchLn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Feb 10 22:55:36 1997; Maspar time 1092.42 Seconds
Tabular output not generated. 916.033 Million cell updates/sec

Title: >US-08-674-311-1
Description: (1-2269) from US08674311.seq (1 of 2)
Perfect Score: 2269
N.A. Sequence: 1 GAATTCGCGTCGACGCTGCT.....AAAAAAGCGGATTC 2269
Comp: CTTAAGCGGAGCGGTCGACGA.....TTTTTTTTCGCTTAAG

Scoring table:
TABLE default
Gap 6

Match STD : Dbase 0; Query 0

Searched: 630489 seqs, 220513910 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99
EST-SYS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174
175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
195:EST195 196:EST196 197:EST197 198:EST198 199:EST199

Statistics: Mean 12.257; Variance 2.969; scale 4.129
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred.	No.
1	1	67	3.0	168	50	HUMNR726		Human keratinocyte CD	2.35e-57	
	2	49	2.2	400	134	G17388		human STS SHGC-14428	8.74e-32	
	3	49	2.2	400	173	HS388293		human STS SHGC-14428	8.74e-32	
	4	41	1.8	473	60	N45983		YY57B12.s1 Homo sapie	2.97e-21	
	5	33	1.5	344	45	HS1MBB133		H.sapiens partial mRN	1.53e-11	
	6	27	1.2	284	48	HUM231B05B		Human aorta CDNA 5'-e	4.76e-05	
	7	28	1.2	285	51	MG5EG113		M.musculus DLA-113 mr	4.53e-05	
	8	27	1.2	431	86	R50764		YG70C04.s1 Homo sapie	4.76e-05	
	9	28	1.2	452	55	N30827		YW70A07.s1 Homo sapie	4.53e-05	
	10	27	1.2	490	168	HSW03788		za61g02.r1 Soares fet	4.76e-05	
	11	27	1.2	490	147	W03788		za61g02.r1 Soares fet	4.76e-05	
	12	27	1.2	492	132	G08836		human STS CHLC.ATRA21A	4.76e-05	
	13	27	1.2	565	60	N45982		YY57B12.r1 Homo sapie	4.76e-05	
	14	25	1.1	144	87	R56208		YG90A03.s1 Homo sapie	4.29e-03	
	15	25	1.1	186	166	HS933319		zb57e06.s1 Homo sapie	4.29e-03	
	16	26	1.1	228	147	W04550		zb70D01.s1 Homo sapie	4.68e-04	
	17	26	1.1	228	163	HS550335		zb91e04.s1 Soares par	4.68e-04	
	18	26	1.1	254	132	G08783		human STS CHLC.ATRA5H0	4.68e-04	
	19	26	1.1	263	161	HS362315		zb70D01.s1 Homo sapie	4.68e-04	
	20	26	1.1	263	143	N93362		zb70D01.s1 Homo sapie	4.68e-04	
	21	25	1.1	290	49	HUM337610B		Human aorta CDNA 5'-e	4.29e-03	
	22	25	1.1	305	48	HUM252H12B		Human aorta CDNA 5'-e	4.29e-03	
	23	25	1.1	306	132	G09779		human STS CHLC.ATRA24E	4.29e-03	
	24	25	1.1	306	55	N29618		Yw67D03.s1 Homo sapie	4.29e-03	
	25	26	1.1	317	7	H03560		YJ37F04.r1 Homo sapie	4.68e-04	
	26	26	1.1	317	7	N36600		YJ37F04.r1 Homo sapie	4.68e-04	
	27	26	1.1	329	57	N36600		YX88A03.r1 Homo sapie	4.29e-03	
	28	26	1.1	330	84	R45024		YX36H09.s1 Homo sapie	4.68e-04	
	29	26	1.1	332	62	NS2043		Y208H05.s1 Homo sapie	4.68e-04	
	30	25	1.1	342	61	N47304		YY74A07.r1 Homo sapie	4.29e-03	
	31	26	1.1	347	135	HS217AC3		H. sapiens (DIS436) D	4.68e-04	
	32	26	1.1	359	46	HUM020E05A		Human fetal brain cdn	4.68e-04	
	33	26	1.1	380	174	HSB334ZAS		H.sapiens (D232301) D	4.68e-04	
	34	26	1.1	380	136	HSB334ZAS		H.sapiens (D232301) D	4.68e-04	
	35	26	1.1	396	161	HS413323		za73C06.r1 Soares fet	4.68e-04	
	36	25	1.1	396	146	W04113		za73C06.r1 Soares fet	4.68e-04	
	37	25	1.1	402	46	HUM065F03A		Human fetal brain cdn	4.29e-03	
	38	26	1.1	412	12	H19304		Ym44A06.s1 Homo sapie	4.29e-03	
	39	25	1.1	421	62	N51355		Yz16A05.s1 Homo sapie	4.68e-04	
	40	25	1.1	439	9	H08801		Y188A03.s1 Homo sapie	4.29e-03	
	41	25	1.1	452	53	N24083		YX67A07.s1 Homo sapie	4.29e-03	
	42	25	1.1	471	83	R42451		YQ20A02.s1 Homo sapie	4.29e-03	
	43	25	1.1	472	143	N92751		zb26A07.s1 Homo sapie	4.29e-03	
	44	25	1.1	472	165	HS751318		zb26A07.s1 Homo sapie	4.29e-03	
	45	26	1.1	533	73	R06864		Yf11e10.s1 Homo sapie	4.68e-04	
45	26	1.1	685	153	W232325		65E9 Human retina cDN	4.68e-04		
ALIGNMENTS										
RESULT	1									
LOCUS			HUMNR726	168 bp	mrna			EST	01-DEC-1994	
DEFINITION			Human keratinocyte CDNA, clone 726.							
ACCESSION			D29573							
NID			G959506							
KEYWORDS			EST(expressed sequence tag).							
SOURCE			Homo sapiens Epidermis Keratinocyte CDNA to mRNA.							
ORGANISM			Homo sapiens							
			Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.							
REFERENCE			1 (bases 1 to 168)							

ALIGNMENTS

LOCUS	1	HUMNR726	168 bp	mRNA	EST	01-DEC-1994
DEFINITION	1	Human keratinocyte CDNA, clone 726.				
ACCESSION		D29573				
KEYWORDS		9599506				
SOURCE		Homo sapiens				
ORGANISM		Homo sapiens				
REFERENCE		1 (bases 1 to 168)				

AUTHORS Konishi, K., Morishima, Y., Ueda, E., Nonomura, K., Kibe, S., Yamashita, K., and Yasuno, H.
JOURNAL Unpublished (1994)
REFERENCE 2 (sites)
AUTHORS Konishi, K., Morishima, Y., Ueda, E., Kibe, Y., Nonomura, K., Yamashita, K., and Yasuno, H.
TITLE Cataloging of the genes expressed in human keratinocytes: analysis of 607 randomly isolated cDNA sequences
JOURNAL Biochem. Biophys. Res. Commun. 202 (2), 976-983 (1994)
MEDLINE 94374994
COMMENT Submitted (11-Mar-1994) to DDBJ by:
 Kiyofumi Yamashita
 Department of Dermatology
 Kyoto Prefectural University of Medicine
 Kamigyo-Ku
 Kyoto, Kyoto, 602
 Japan
 Phone: 75-251-5587
 Fax: 75-251-5587
 PROJECT="human.epidermal_keratinocyte_ESTs",
 Location/Qualifiers
 1..168
 /organism="Homo sapiens"
 /cell_line="Primary culture"
 /cell_type="Keratinocyte"
 /sequenced_mol="cDNA to mRNA"
 /tissue_type="Epidermis"

BASE COUNT 43 a 26 c 37 g 62 t
ORIGIN

Query Match 3.08; Score 67; DB 50; Length 168;
 Best Local Similarity 93.38; Pred. No. 2,35e-57;
 Matches 83; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Db 19 agattggaataat-ggtgacacagcctgtagatccagaataattagaaggaactt 77
 |||||||
 Oy 153 AGATTGGAATATATGCTGGAACAGCGCTGATGATCCAGAAATTTAGAGAGAACTG 212
 |||||||

Db 78 ttaaatatgtgatagac-cacattgagcaag 105
 |||||||
 Oy 213 AAAAATATGTCGATATCCATTTGGCAG 241
 |||||||

RESULT 2
LOCUS G17388 400 bp DNA STS 05-MAR-1996
DEFINITION human STS SHGC-14428 clone PG-7320.
ACCESSION G17388
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human plasmid clones, generated from a lymphoblastoid cell line from a human male. Localized to human chromosome 9 by analysis on the NIGMS Human/Rodent Somatic Cell Hybrid Panel #1, Coriell Institute for Medical Research, Camden, NJ 08103.
ORGANISM Homo sapiens
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 400)
AUTHORS Myers, R.M.
JOURNAL Unpublished (1996)
COMMENT Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myerseshgc.stanford.edu
 Primer A: TGTGATTAATCCATTGGCA
 Primer B: TCACAAATCAAGCAGCAG
 STS size: 90

PCR Profile:
 Initial incubation: 94 degrees C for 90 seconds
 Denaturation: 94 degrees C for 15 seconds
 Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600

Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Tag Polymerase: 0.05 units/uL
 Total Vol: 10 uL

Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.3

FEATURES
 source
 1..400
 /organism="Homo sapiens"
 /note="human"
 61..150
 /map="9"
 61..80
 /map="9"
 primer_bind
 complement(131..150)
 primer_bind
 /map="9"

BASE COUNT 95 a 74 c 83 g 101 t 47 others
ORIGIN

Query Match 2.28; Score 49; DB 134; Length 400;
 Best Local Similarity 98.28; Pred. No. 8.74e-32;
 Matches 55; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 28 atccagaattttagaaggaaga-ctgaataatcgtgatactcattggaag 82
 |||||||
 Oy 186 ATCCAGAAATTTAGAGAGAGAACTGAATAATATGTCGATCTCATTTGGCAG 241
 |||||||

RESULT 3
ID H5388293 standard; DNA; STS; 400 BP.
AC G17388;
DT 07-MAR-1996 (Rel. 47, Created)
DR 07-MAR-1996 (Rel. 47, Last updated, Version 1)
DE human STS SHGC-14428 clone PG-7320.
KW primer; sequence tagged site; STS sequence.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
RN [1]
RP 1-400
RA Myers R.M.;
RT ;
RL Unpublished.
CC Contact: Richard M. Myers Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine Department of Genetics
 M-344, Stanford, CA 94305, USA Tel: 4157259687 Fax: 4157259689
CC Email: myerseshgc.stanford.edu
CC Primer A: TGTGATTAATCCATTGGCA
CC Primer B: TCACAAATCAAGCAGCAG
CC STS size: 90 PCR Profile: Initial C
 incubation: 94 degrees C for 90 seconds Denaturation: 94 degrees C
 for 15 seconds Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds PCR Cycles: 30 Thermal
 Cycler: Perkin Elmer 9600 Protocol: Template: 25 ng Primer: each 1
 uM dNTPs: each 200 uM Tag Polymerase: 0.05 units/uL Total Vol: 10
 uL Buffer: MgCl2: 2.5 mM KCl: 50 mM Tris-HCl: 20 mM pH: 8.3
CC Chromosome 9, NCBI gi: 1214814
CC Location/Qualifiers
FF key

FT	source	1..400	/organism="Homo sapiens"
FT			/note="human"
FT	STS	61..150	
FT		/map="9"	
FT	primer_bind	61..80	
FT		/map="9"	
FT	primer_bind	complement(131..150)	
FT		/map="9"	
FT	Sequence	400 bp; 95 A; 74 C; 83 G; 101 T; 47 other;	
Query Match	2.2%; Score 49; DB 173; Length 400;		
Best Local Similarity	98.2%; Pred. No. 8,74e-32;		
Matches	55; Conservative	0; Mismatches	0; Indels 1; Gaps 1;
Db	28 atccagaatttagaagaaga-citgaataatgtgatactccattggcaag	82	
Qy	186 ATCCGAAATTTTGAAGGAGAGACTGAAATAATGTGATCTCATTTGGCAG	241	
RESULT	4		
LOCUS	M45993	473 bp	MRNA
DEFINITION	YY57b12.s1 Homo sapiens CDNA clone 277631 3'.	EST	14-FEB-1996
ACCESSION	M45983		
NID	g1187149		
KEYWORDS	EST.		
SOURCE	human clone=277631 primer-m13 -40 forward library=Soares multiple sclerosis 2bdkhsp vector=pt730 (pharmacia) with a modified polylinker V-type: phagemid host=DH10B (ampicillin resistant) Nsite1=Not I Nsite2=Eco RI 46 year old male. 1st strand CDNA was primed with a Not I - oligo(dG) primer [5'-GGTACCACATCTGAGAGGGAGGGCGGCATTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T vector (pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Felima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).		
ORGANISM	Homo sapiens		
	Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (Phases 1 to 473)		
REFERENCE	Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,J., Hulman,M., Kucada,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaszki,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.		
TITLE	The Washu-Merck EST Project		
JOURNAL	Unpublished (1995)		
COMMENT	Contact: Wilson RK Washu-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu High quality sequence stops: 399 Source: IMAGE Consortium, LNL. This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Location/Qualifiers 1..473 /organism="Homo sapiens" /clone="277631" /note="human"		
FEATURES			
source			
BASE COUNT	187 a 69 c 80 g 137 t		
ORIGIN	<1..>473		

Query Match	1.8%	Score 41;	DB 60;	Length 473;
Best Local Similarity	79.28;	Pred. No. 2.97e-21;		
Matches	76;	Conservative	0;	Mismatches 17; Indels 3; Gaps 2;
Db	2	gaatattcaaaacctttatttcgtgtaagaataagaccattt--gttagattac-a 58		
Oy	905	GAATGTCGAGAAACCTCCATTACCTGAGAGAAATATGGCCAGATTTCGTTTATTTCACA 964		
Db	59	agacattaaaataacatgctgcgccagagaaaga 94		
Oy	965	AGACATTAAAGTAGCATGCTGCCAGAGAAAGA 1000		
RESULT	5	HSIMB13	344 bp	RNA
LOCUS				EST
DEFINITION		H.sapiens partial mRNA; expressed sequence tag (clone lmbd-est 133).		
ACCESSION		X93833		
NID		91103448		
KEYWORDS		EST; expressed sequence tag.		
SOURCE		human.		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (phases 1 to 344)		
AUTHORS		Moschonas,N.K., Argyrokastilis,A. and Vrontou,S.		
REFERENCE		Unpublished		
JOURNAL		2 (bases 1 to 344)		
REFERENCE		Moschonas,N.K.		
AUTHORS		Moschonas,N.K.		
TITLE		Direct Submission		
JOURNAL		Submitted (23-NOV-1995) Moschonas N.K., Institute of Molecular Biology and Biotechnology, FORTH, P.O. Box 1527, Heraklion 711 10, Crete, Greece		
COMMENT		K562-cDNA library normalization: Puzirev et al., Molecular Biology (Molecularyarnaya Biologiya), 29:58-61(1995).		
FEATURES		both strands read (asymmetric PCR sequencing)		
SOURCE		no significant homology to GenBank entries.		
LOCATION/Qualifiers		1..344		
ORGANISM="Homo sapiens"		/cell_line="K562 (erythro-leukemic)"		
/dev_stage="embryonal/fetal with respect to globin expression"		/clone_lib="normalized K562 cDNA library"		
/clone="lmbd-est 133"		<1..>344		
BASE COUNT		96 a	89 c	72 g 87 t
ORIGIN		mrna		
Query Match	1.5%	Score 33;	DB 45;	Length 344;
Best Local Similarity	77.0%;	Pred. No. 1.53e-11;		
Matches	47;	Conservative	0;	Mismatches 14; Indels 0; Gaps 0;
Db	253	atccataataagcaagagctcttcctctcttccacattactctctgtaaggaagt 312		
Oy	1353	ATCCATTAATAATACCAAGAGCTCCATTCTCTCCACATCTTTGGCGAGTATAGAAAGT 1294		
Db	313	a 313		
Oy	1293	A 1293		
RESULT	6	HUM21B05B	284 bp	mrna
LOCUS				EST
DEFINITION		Human aorta cDNA 5'-end GEN-231B05.		
ACCESSION		D62007		
NID		9965783		
KEYWORDS		EST(expressed sequence tag); Human aorta;		
SOURCE		similar to none(May 29,1995).		
ORGANISM		Homo sapiens (library: Clontech human aorta polyA+ mRNA (#6572))		
FEATURES		cDNA to mRNA.		
LOCATION/Qualifiers		Homo sapiens		
ORGANISM		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;		

/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5'-AATCGAAGATTAATTAAGACCTTTTCTTTTCTTTT 3'] (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo." /clone="297074" /clone_lib="Soares fetal liver spleen INFLS" /sex="male" /dev_stage="20 week-post conception fetus" /lab_host="DH10B (ampicillin resistant)" <1..>490 91 c 64 g 186 t

Query Match 1.28; Score 27; DB 147; Length 490; Best Local Similarity 84.68; Pred. No. 4,76e-05; Matches 33; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 319 ctttttttttttttaagaatgacatccatttatt 357
|||||
Cp 2261 CTTTTTTTTTTTTTAAAGACGCTTCACCTTTATT 2223

RESULT 12
LOCUS G08836 492 bp DNA STS 08-AUG-1995
DEFINITION human STS CHLC.ATA21A09. P15975 clone ATA21A09.
ACCESSION G08836
NID 9839386
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human vector-pCCPI host-E.coli dut-ung+ (DH10B) Marker Selected genomic DNA prepared from XY individual of French nationality.

ORGANISM

Homo Sapiens; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 492)
Murray, J., Sheffield, V., Weber, J. L., Duyk, G. and Buetow, K. H.
Cooperative Human Linkage Center
Unpublished (1995)
Synonyms: ATA21A09, CHLC.ATA21A09. #15719
Contact: Dr. Jeffrey C. Murray
Institution: The University of Iowa
Department of Pediatrics, Iowa City, IA 52242, USA
Tel: (319) 356-3508
Fax: (319) 356-3347
Email: jeff-murray@uiowa.edu

COMMENT

Primer A: ATGCAATCCAGAGCTTCAC
Primer B: ATGATATATGCTCGGAGA
STS size: 337
PCR Profile:

denature: 30 seconds at 94 degrees C
annealing: 75 seconds at 55 degrees C
extension: 15 seconds at 72 degrees C
PCR cycles: 27
extension: 6 minutes at 72 degrees C

Protocol:

Template: 30ng genomic DNA
Primer: each 1.5 pmole
DNTPs: each 200 uM
Taq Polymerase: 0.3 units
Total Vol: 10 uL

Buffer:

MgCl2: 1.5mM
KCl: 50mM
Tris: 10mM

FEATURES
source Location/Qualifiers
1..492
/organism="Homo Sapiens"
/note="human"

STS
primer_bind 71..407
primer_bind 71..90
BASE COUNT 231 a 76 c 75 g 103 t 7 others
ORIGIN

Query Match 1.28; Score 27; DB 132; Length 492; Best Local Similarity 73.68; Pred. No. 4,76e-05; Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db 270 taataataataataataagccctgaanaacntgcaccccttaca 322
|||||
Oy 1519 TAATATGATAGTATTAATTAATTAATTAATGAACATCTGAATCCCTTGCAA 1571

RESULT 13
LOCUS N45982 565 bp mRNA EST 14-FEB-1996
DEFINITION Y57b12.r1 Homo sapiens cDNA clone 277631 5'.
ACCESSION N45982
NID 9187148
KEYWORDS EST.
SOURCE human clone-277631 primer-T7 library-Soares multiple sclerosis 2NbhMSP vector-pT73D (Pharmacia) with a modified polylinker V-type: phagemid host-DH10B (ampicillin resistant) Rs1et-Not I Rs1et2-Eco RI 46 year old male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTACCAATCTGAACGACGCGCCGACATTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).

ORGANISM

Homo Sapiens; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 565)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence stops: 409
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lln.gov) for further information.

FEATURES
source Location/Qualifiers
1..565
/organism="Homo sapiens"
/clone="277631"
/note="human"

BASE COUNT 194 a 80 c 100 g 188 t 3 others
ORIGIN

THIS PAGE LEFT BLANK

(TM)

Distribution rights by Intelligenetics, Inc.

```

Mpsrch_nn      n.a. - n.a. database search, using Smith-Waterman algorithm

```

```
Run on: Mon Feb 10 23:44:01 1997; Maspar time 616.28 Seconds
```

tabular output not generated;

Description: (122-970) from US08674311.seq (2 of 2)

N.A. Sequence: 122 ATGGCCTCTGGCACCACCAC.....CTGTTTATTACCAAGACAT 970

Comp:

Scoring table: TABLE default

Searched: 279077 seqs, 411808665 bases x 2

Post-processing: Minimum Match 0%

Database:

10:PRO1 11:PRO2 12:ROD 13:SYN 14:UNC 15:VIR

base:

15:BCV1 17:BC12 18:BC13 19:BC14 20:BC15 21:BC16 22:BC17
23:BC18 24:BC19 25:INV1 26:INV2 27:INV3 28:INV4 29:INV5
30:INV6 31:INV7 32:MAM1 33:MAM2 34:MAM3 35:VRT1 36:VRT2
37:VRT3 38:PAT1 39:PAT2 40:PAT3 41:PHG 42:PLN1 43:PLN2
44:PLN3 45:PLN4 46:PLN5 47:PLN6 48:PLN7 49:PLN8 50:PR11
51:PR12 52:PR13 53:PR14 54:PR15 55:PR16 56:PR17 57:PR18
58:PR19 59:PR10 60:PR11 61:PR12 62:PR13 63:ROD1
64:ROD2 65:ROD3 66:ROD4 67:ROD5 68:ROD6 69:ROD7 70:ROD8
71:STR 72:SYN 73:DNA 74:VRL1 75:VRL2 76:VRL3 77:VRL4
78:VRL5 79:VRL6 80:VRL7 81:VRL8

Database:

89:PLN 90:PRI 91:ROD 92:STR 93:SYN 94:UNA 95:VRL

Database: u-emb146_94

Statistics: Mean 11.029; Variance 4.043; scale 2.728

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description	Pred. No
1	849	100.0	2269	HSU22233	Human methylthioadenosine	0.00e+00
2	35	4.1	14071	SPAC1659	S.pombe chromosome I	8.94e-10
3	31	3.7	38946	CEB0228	Caenorhabditis elegans	1.52e-06

4	27	3.2	5550	62	RCTYBC1	R. rubrum cytochrome	1.63e-033
c	5	2.8	3205	60	HUNNEKB34	Human factor KBFI MRN	2.03e-011
c	6	2.8	3625	60	HUNNEKB	Human nuclear factor	2.03e-011
c	7	2.8	9155	54	HSNFX11.12	H. sapiens NFKAI gene	2.03e-011
c	8	2.8	29098	26	CEJF18C5	Caenorhabditis elegans	9.33e-010
c	9	2.6	42582	22	CEB0361	Caenorhabditis elegans	4.06e+000
c	10	2.6	819	19	HEAH10655	Haemophilus influenza	4.06e+000
c	11	2.6	913	66	MUSG1CYC	Mus musculus D-type G	4.06e+000
c	12	2.6	12500	82	HIUJ2694	Haemophilus influenza	4.06e+000
c	13	2.6	12500	10	HIJ32694	Haemophilus influenza	4.06e+000
c	14	2.6	12500	20	HIJ32694	Haemophilus influenza	4.06e+000
c	15	2.6	41453	26	CER06F6	Caenorhabditis elegans	4.06e+000
c	16	2.6	42159	83	CEB0024	Caenorhabditis elegans	4.06e+000
c	17	2.6	42154	3	CEB0024	Caenorhabditis elegans	4.06e+000
c	18	2.6	100000	10	HIU00075	Haemophilus influenza	4.06e+000
c	19	2.2	100000	20	HIU00075	Haemophilus influenza	4.06e+000
c	20	2.5	288	37	S78512	myoD-transcriptional	1.67e+001
c	21	2.5	292	45	MSND0U.	Alfalfa nodulin mRNA	1.67e+001
c	22	2.5	870	65	MMU35374	Mus musculus purine n	1.67e+001
c	23	2.5	1153	17	CHEAL165	C. thermaceticus 16S r	1.67e+001
c	24	2.5	1338	64	MPMPMR	M. musculus Np-b mRNA	1.67e+001
c	25	2.5	1890	17	DVPR15	D. vulgaris gene for p	1.67e+001
c	26	2.5	2049	30	PFASPL	Plasmodium berghei for p	1.67e+001
c	27	2.5	2263	32	BTCRG	Bovine mRNA for carli	1.67e+001
c	28	2.5	2458	16	ACERSRTA	A. calcocecticus esta	1.67e+001
c	29	2.5	2585	21	NGIRXAGNA	S. typhimurium ttrX ge	1.67e+001
c	30	2.5	4107	12	MMU20238	Mus musculus grase-a	1.67e+001
c	31	2.5	4107	65	MMU20238	Mus musculus grase-a	1.67e+001
c	32	2.5	5426	59	HUMK1AA08	Human mRNA for KIA01	1.67e+001
c	33	2.5	6577	50	CALMYCG	Callithrix jacchus MY	1.67e+001
c	34	2.5	8124	22	PMUNC2	Prodesophila genes unci	1.67e+001
c	35	2.5	11392	28	DPRGAT	Drosophila pseudobsc	1.67e+001
c	36	2.5	30570	52	HS91K3A	Human DNA sequence fr	1.67e+001
c	37	2.5	33950	25	CE37B12	Caenorhabditis elegans	1.67e+001
c	38	2.5	37172	2	SPAC22E12	S. pombe chromosome I	1.67e+001
c	39	2.5	37172	48	SPAC22E12	S. pombe chromosome I	1.67e+001
c	40	2.5	38383	25	CE3F5H8	Caenorhabditis elegans	1.67e+001
c	41	2.5	40592	54	HS191P1	Human DNA sequence fr	1.67e+001
c	42	2.5	43410	83	CEC01G6	Caenorhabditis elegans	1.67e+001
c	43	2.5	43410	3	CEC01G6	Caenorhabditis elegans	1.67e+001
c	44	2.5	119704	45	OSCHLPLX	O. siemensis chloroplast	1.67e+001
c	45	2.5	119704	5	OSCHLPLX	O. siemensis chloroplast	1.67e+001

ALIGNMENTS

RESULT	1			
LOCUS	HSU22233	2269 bp	mRNA	PRI
DEFINITION	Human methylthioadenosine phosphorylase (MTAP)		25-NOV-1995	
ACCESSION	U22233		mRNA, complete	cds
NID	9847723			
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			

REFERENCE
AUTHORS

TITLE

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 92 (14), 6489-6493 (1995)

REFERENCE 2 (bases 1 to 2269)

AUTHORS

JOURNAL

FEATURES

```

source
1..2269
/clone="18-11 CDNA"
/map="9p21"
/chromosome="9"
/organism="Homo sapiens"
/sex="male"
/cell_line="primary culture"
/cell_type="fibroblast"
/tissue_type="epidermis"
122..973
/clone="MTAP"
/codon_start=1
/product="methylthioadenosine phosphorylase"
/db_xref="pid:9847724"
/translation="MAGSTTTAVKIGIIGTGLDDELLEGRTEKYNDTPGRKPSD
LIIKIKNVDCIILARHROHTIMPSVNVNQAATIMAKREGCHVITVTAAGSTIEET
OPGDIIVIDIDRTMROSFYDGSRCARICHIPMAEPKPTREVLITAKLGL
LRCHSKCTMYTIEGPRSSRAESPMFTWADYINMTYPEVYLAKEGICTASTAMA
TDYDKWEHERAVSYDRVLKLIKENAKAKSLTLTTIPIQISTEWSLTLNKLNAOF
SVLLEPRH"

BASE COUNT      725 a      407 c      490 g      647 t

Query Match      100.0%; Score 849; DB 56; Length 2269;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 849; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 122 atgacctctggccaccacccacccgctgaagattggaataattgtggaagagcctg 181
122 ATGGCCTTGCCGCCACCCACCCCGTGAAGATTGGAATAATTGTTGACAGCGCTG 181
182 gatgatccagaagaattttgaagaagaactaaaaattgtggatactcattggcaag 241
182 GATGATCCAGAAATTTTGAAGAGAACTAAAAAATGTGGATCTCCATTGGCAG 241
182 gatgatccagaagaattttgaagaagaactaaaaattgtggatactcattggcaag 241
182 GATGATCCAGAAATTTTGAAGAGAACTAAAAAATGTGGATCTCCATTGGCAG 241
242 ccactcgatgccttaattttggggaagataaaaaattgtgcatcctccttgaag 301
242 CCATCTGATGCTTAAATTTGGGGAAGATAAAAAATGTGATTCATCTCTGCAAG 301
242 CCATCTGATGCTTAAATTTGGGGAAGATAAAAAATGTGATTCATCTCTGCAAG 301
302 catggaagcagcacaccatcatctgctcctcaaaaggtcaactaccagcgcaacatctg 361
302 CATGGAAGCAGCACACCATCATCTGCTTCAAAAGGTCAACACACGAGCAATCTGGCT 361
302 CATGGAAGCAGCACACCATCATCTGCTTCAAAAGGTCAACACACGAGCAATCTGGCT 361
362 ttgaagaagaagggctgtacacatgcatgtagcccaagctgtgctcctctgaagag 421
362 TTGAAGAAGAAGGGCTGTACACATGTCATAGTACACAGCTTGTGCTCTTGAAGAG 421
362 TTGAAGAAGAAGGGCTGTACACATGTCATAGTACACAGCTTGTGCTCTTGAAGAG 421
422 gagatccagccggcggaatgtcatattatgataagttcatitgaagagaccactatgaga 481
422 GAGATTCAGCCCGGCGAATGTCTATATGATCAGTTCATGACAGGACCACTATGAGA 481
482 cctcagctctcatatgatgagaatcatcttgcgcagagagtggtgcatatccaatg 541
482 CCTCACTCTCATATGATGAGAATCATCTTGCCTGCGAGAGTGTGCCATATTCGAAG 541
482 CCTCACTCTCATATGATGAGAATCATCTTGCCTGCGAGAGTGTGCCATATTCGAAG 541
542 gctgagacgcttttgcgcccaaaacgagagaggttcttatagaagactgtaagaagc 601
542 GCTGAGACGCTTTTGCGCCAAAACGAGAGAGGTCTTATAAGAGACTGTAAGAAGC 601
542 GCTGAGACGCTTTTGCGCCAAAACGAGAGAGGTCTTATAAGAGACTGTAAGAAGC 601
602 ctcccggtgacatcaaaaggggacaaatgttgcacaaatgagaggaactggtttta 661
602 CTCCGGTGCATCAAAAGGGGACAAATGTGTCAATCAAGGAGACTGCTTTTAACTCC 661
602 CTCCGGTGCATCAAAAGGGGACAAATGTGTCAATCAAGGAGACTGCTTTTAACTCC 661
662 gcaagaagaattcatgttcgcacactggggggcgagatgttatcaaatgacacagttcca 721
662 GCAAGAAGATTCTATGTCTCCGACCTGGGGGCGAGATGTATCAACATGACCAAGTTCCA 721
662 GCAAGAAGATTCTATGTCTCCGACCTGGGGGCGAGATGTATCAACATGACCAAGTTCCA 721
722 gaggtgtgtcttctgtaagaagcttgaattgtgtacgaagtaatcgccatgagcagat 781
722 GAGGTGTGTCTTCTGTAAGAAGCTTGAATTGTGTACGAAGTAATCGCCATGAGCAGAT 781
722 GAGGTGTGTCTTCTGTAAGAAGCTTGAATTGTGTACGAAGTAATCGCCATGAGCAGAT 781
782 tatgactgttgaagaagacagaagaagcagtttcggttgaccgggtcttaagaccctg 841
782 TATGACTGTTGGAAGACAGACGAGAGAGTTCGTTGGTGCACCGGCTCTTAAGACCTG 841

```

```

QY 782 TATGACTGTTGGAAGACAGACGAGAGAGTTCGTTGGTGCACCGGCTCTTAAGACCTG 841
Db 842 aaagaagaagcctaataaagccaaagcttactgtcactacatcacatagaggtcc 901
QY 842 AAAAGAAAACCTATTAACCCAAAAGCTTACTGCTCACTACCAATCCTCAGAAAGGCTCC 901
Db 902 acagaatgctcagaaacctcacaaccatgaagaatatgagccagtttctgttata 961
QY 902 ACAGAATGCTCAGAAACCTCCTATACCTGGAAGATATGCGCCAGTTTCTGTTATTA 961
Db 962 ccaagacat 970
QY 962 CCAGACAT 970

RESULT 2
LOCUS SPAC16C9 14071 bp DNA PLN 22-OCT-1995
DEFINITION S.pombe chromosome I cosmid c16C9.
ACCESSION 254366
NID 91019812
KEYWORDS MAK32 protein homologue; NAM7 protein homologue; NOT4 homologue;
serine/threonine protein kinase; zinc finger protein.
SOURCE fission yeast.
ORGANISM Schizosaccharomyces pombe
Eukaryotes; mitochondrial eukaryotes; Eumycota; Ascomycotina;
Archaeoscomycetes; Schizosaccharomycetales;
Schizosaccharomycetaceae; Schizosaccharomycetes.
1 (bases 1 to 14071)
Badcock, K. and Churcher, C.M.
REFERENCE
JOURNAL
AUTHORS
REFERENCE
TITLE
JOURNAL
COMMENT
Details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, http://www.sanger.ac.uk/
yeastpub/svw/home.html)
computer coding regions (CDS) have been predicted with the help of
protein analysis using the GeneFinder program in Pombase (an ACEDB
database) with additional predictions for the branch-acceptor sites
supplied by the program Sp3splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.
CDS are numbered using the following system eg SPAC5H10.01c. SP (S.
pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
(complementary strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
cosmid c16C9 is overlapped by cosmid c468 at the 5' end and by
cosmid c2611 at the 3' end.
location/Qualifiers
1..14071
/organism="Schizosaccharomyces pombe"
/strain="972h-"
/chromosome="I"
/clone="cosmid c16C9"
/map="II"
1..267
/misc_feature
/note="overlap with 468"
complement(96..101)
/misc_feature
/note="splice donor"
complement(101..179,217..256))

```

/gene="SPAC16c9.01c"
/note="SPAC16c9.01c, partial spliced CDS, len: >39,
similar to SW:MK32_YEAST P23060 MKR32 protein (46.78
identity in 30 aa overlap); pid:e204615"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019813"
/translation="MDYPSKDVETVTGMFIIDIEFDGCKKIDIGAGSGS"
/complement(180..193)
/gene="SPAC16c9.01c"
/note="splice branch and acceptor sequence,
ctaagaatgaag"
/complement(211..216)
/gene="SPAC16c9.01c"
/note="splice donor sequence, gtaaaa"
/complement(672..682)
/note="splice branch and acceptor sequence, ctaaacattag"
/complement(join(777..1037,1137..1697,1747..1848))
/gene="SPAC16c9.02c"
/note="SPAC16c9.02c, unknown, len: 307, similar to
SW:YPE2_RHORO P23139 hypothetical 25.8 kd protein (46.98
identity in 128 aa overlap); pid:e204616"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019814"
/translation="MNKQAEPIILGVIGSGFYDLPGEDYVESVNPITPMYSPIS
IARTSGFLIFLARHGVIHTPTPEVPSANIALKSLGLVATVSFSAVSLIEDIP
PEDVLPQIIDRTLCANPTFEFGCAVHSGFDPDQDYEILSSGSLKNGSKL
HTRKRGDLYVMEGPAFSTRASNLYRWGASIIINNSVPEAKLAREAIAYOMC
MATYDCRMNEEVEVETVEHISNNNDNAKIFLEAVKLEAPLLQGLFRLRES
VONGIOTNHRKNDAIRLOELFPNLITPA"
/complement(1038..1051)
/gene="SPAC16c9.02c"
/note="splice branch and acceptor sequence,
ctaacaataag"
/complement(1131..1136)
/gene="SPAC16c9.02c"
/note="splice donor sequence, gtaagt"
/complement(1698..1713)
/gene="SPAC16c9.02c"
/note="splice branch and acceptor sequence,
ctaagttatcag"
/complement(1741..1746)
/gene="SPAC16c9.02c"
/note="splice donor sequence, gtagta"
3172..4668
/gene="SPAC16c9.03"
/note="SPAC16c9.03, unknown, len: 498, similar to
SW:YHO_YEAST P38861 hypothetical 59.1 kd protein (54.98
identity in 494 aa overlap); pid:e204617"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019815"
/translation="MDMDSGSGVAFOPSEPIASTIICCEGVPPTPPAAMCMDCIK
MTDTIGIPRESTVNHCRCEERYMOPNNMNIAPLESRELMATCIKRLNVRILY
DANITWEPHSRIKRLIVQKEAFNTLIQOSFVERVYNNTCOPDCAIYTHIMK
AVCQKRVLRKRTFLLEQILIKRKAMNTVNIKEKQIDFEGORAHAKIVEL
SAVPIRIKSGSEELISEDFKSNITKTYSEIEVPLCKDLVCLPKVAAHGNIAQ
LVCTKGPTRIFLDPLOTCDMLPSIYRTPPALADIPETFEIYADVLDLGPIN
GKVALADELIKSSDSTHLTRHGLIANGVTVYHVLATNPNNEVYDILREDSI
PEVIVKTYRQTKKNNRMLKTIQMKADRYKODIEREERYELFIQNLBEDPEL
RQGVNLKAPKATAVADTDHDEDEVEDIPQISVDLDDVAMH"
/complement(join(5081..6317,6406..6585,6636..6688))
/gene="SPAC16c9.04c"
/note="SPAC16c9.04c, regulatory protein, len: 489, similar
to SW:NOT4_YEAST P34909 general negative regulator of
transcription subunit 4 (NOT4) (38.88 identity in 417 aa
overlap); pid:e204618"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019816"
/translation="MLKTOEIISSDEDDMCPLCMEIEIDISDNFKPCOGGYVGR
CMHRIKEDLNGRCACRRLTYEENVQMPVYAEEMKMDLHKNNRKRERKEKREVELS

NRKHIANIRVOKNLAVNGLSPKVANEENINILKGEYRGYKILIKANKAAN
SANGHGVITTYORKEDAAIAAIDSGVSDGHLRASITTYTKCYSILNOCPPMS
CMTLHEGDEVDSTIEDLASLQHTRELSPPVNVNATSPSPSPLEFKVPLPVKT
PLEEANSPPAQNOHITTVHVOVSMETSLSTNNATVPAPYSSAAVNVVPGA
TIIHHESSALPTAKMAKLSVQERLRAAVNOQPLDALKSSVTSIPIKOKK
AAKLPESENTKMLKAIINDVLSLSKINFESEGTGFEKQLEMIONLPPLVFNAR
SVIDKEVYPOEKSAAENOPTSLGINNGNVMPPGQS"
/complement(6318..6330)
/gene="SPAC16c9.04c"
/note="splice branch and acceptor sequence, ctaatttttag"
/complement(6400..6405)
/gene="SPAC16c9.04c"
/note="splice donor sequence, gtagtgc"
/complement(6586..6605)
/gene="SPAC16c9.04c"
/note="splice branch and acceptor sequence,
ttaactgaattcacaag"
/complement(6630..6635)
/gene="SPAC16c9.04c"
/note="splice donor sequence, gtagtgc"
7638..8852
/gene="SPAC16c9.05"
/note="SPAC16c9.05, unknown, len: 404, possible zinc
finger protein, similarity to SPAC2f7.07c (in embi entry
250142) 39.18 identity in 115 aa overlap; pid:e204619"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019817"
/translation="MASSINNSQPTVPSISNNSHGDSPVNEGPPSNFKNNLSSTH
SSTHVVNLISODKENDISPVKOKASNSKPNAPLOKSGSLKYLEESQ
TPVRRRRRTIRIVDCSAGRGLEFICCGCPSPCLTIPENITPGSGMC
VTCISKSHHPKHPRLSISOLYMDQNSQVRLPDVLYHTRGIRCTGAKYTE
GEMDTDFSLPTGSSITNLAYCGYCSKPSMGACWYGCQDCOTFHKNEKHAHKS
HDSIGKGMVPRKNAVIRPVLDTISNLRVMSGWQFLMGEPSPDELTFPPL
PVSLTYVSESDGILKDFLVAIGIEAKFRNRRKRELEVIPVKSALPLARHPNLP
IALRLEFKART"
/complement(9398..12205)
/gene="SPAC16c9.06c"
/note="SPAC16c9.06c, len: 935, similar to SW:NAM7_YEAST
P30771 NAM7 protein (UPF1 protein) (51.0% identity in 931
aa overlap); contains PS00017 ATP/GTP-binding site motif A
(P-loop); pid:e204620"
/product="unknown"
/complement(10913..10936)
/gene="SPAC16c9.06c"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
13084..214071
/gene="SPAC16c9.07"
/note="SPAC16c9.07, len: > 329, partial orf, overlaps
cosmid c2611 orf SPAC2611.01, putative serine/threonine
protein kinase; pid:e204621"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019818"
/translation="MGLISTSELPVQESKSSAPVSNFLSPNNLTQTCSPRARAS
TFKEPVLLSQROISHSKSLITELISVKNPKNPQDEEKSASVIEKSKDEHLAK
KGLHNFDLKLEIFLNKSLDLDGQDSLDNDOSGLDNPVLSQNRRIYALEKLAN
SSKEVEYIPVENHNPUNKLSKRTLAETIINDLDEPTGCTVYPMKAKRTLSSETK
KEYQOQKFECSOFAPSRSSIVGPTLIGNSRADDTARMASRMKPSFNNIQQS
SYGASOSTLTSORNDHDKDLNFSYKSLPIPLNNRYSMSSELVSGSGSTTP"
13963..14071
/gene="SPAC16c9.07"
/note="overlap with cosmid c2611"
BASE COUNT 4537 a 2552 c 2494 g 4488 t
ORIGIN

Query Match 4.1%; Score 35; DB 48; Length 14071;
Best Local Similarity 62.1%; Pred. No. 8.94e-10;
Matches 90; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Db 1136 ctgtagccatgcaacaccttgtaggcaatcgcagctgagcttgctcg 1195
||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Cp 779 CTGTGCGCATGGCGATCTTGGCGTACAAATTCACGCTCTCTAGCAAGACACACTGTG 720

Db	1196	gattagctgcacgttggataaacctggccccccagcttcgtatataatcgtttctcgaac	1255
Cp	719	GAAGTGGGCACTGTGATTAACATCCGCCCCCCCAGTGCGGAACAATGAAGCTTTTCGCC	660
Db	1256	ggctactaatgcagaccctcat	1280
Cp	659	GCGAGCTAATAACGACGCTCCCTCGAT	635

LOCUS	3	CELB0228	38946 bp	DNA	INV	30-MAR-1995
DEFINITION		Caenorhabditis elegans cosmid B0228.				
ACCESSION		U23168				
NID		9726361				
KEYWORDS						
SOURCE ORGANISM		Caenorhabditis elegans strain-Bristol N2.				
REFERENCE AUTHORS		Eukaryotae; mitochondrial eukaryotes; eukaryote crown group; Metazoa/Eumyozoa group; Metazoa; Eumetazoa; Bilateria; Pseudocoelomata; Nematoda; Secernentea; Rhabdita; Rhabditidae; Rhabditina; Rhabditoidea; Rhabditidae; Caenorhabditis. 1 (bases 1 to 38946) Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copest,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fulton,L., Gardner,A., Green,P., Hawkins,I., Hillier,L., Jler,M., Johnston,L., Jones,M., Karshay,J., Kilsen,J., Laister,N., Larellalle,P., Lightning,U., Lloyd,C., McKurray,A., Mortimore,B., O'Callaghan,M., Parsons,J., Percy,C., Rilken,L., Roopra,A., Saunders,R., Shownkeen,K., Skilton,N., Smith,A., Sonnhammer,E., Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M., Vaughan,K., Waterston,R., Watson,A., Weinstock,L., Wilkison,Sproat,J. and Woulaman,P. 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans 368 (6466), 32-38 (1994) Nature 368 (6466), 32-38 (1994) 2 (bases 1 to 38946) Leimbach,D. The sequence of C. elegans cosmid B0228 Unpublished (1994) 3 (bases 1 to 38946) Waterston,R. Direct Submission Submitted (21-MAR-1995) Robert Waterston				
JOURNAL MEDLINE REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT		Submitted by: Genome Sequencing Center Department of Genetics, Washington University, St. Louis, MO 63110 USA, and Sanger Centre, Hinxton Hall Cambridge CB10 1RQ, England e-mail: tw@ematoide.wustl.edu and jess@sanger.ac.uk				

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is B0495, 8001 bp overlap; 3' cosmid is C06A8, 750 bp overlap. Actual start of this cosmid is at base position 4495 of CELB0228; actual end is at 6105 of CELC06A8

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation).

Location/Qualifiers

1..38946

/organism="Caenorhabditis elegans"

/strain="Bristol N2"

join(56...502,1135..1583,1735...2017,2219...2342,2390...2842,

CDS

[illegible]

[illegible]

	Query Match	Score 27:	DB 22:	length 5550:
	Best Local Similarity	59.2%:	Pred. No.	1,61e-03:
	Matches	87:	Mismatches	60:
			Indels	0:
			Gaps	
Dd	4267	cgcctcgatgataattaccggcgaacaatatgaaagccctagaaggagggcgtgacacggg	4326	
OY	326	CTTTAAAGGTAACACTACACGAGCAACATCTGGCGTTCCTTGAAGAAGAGGGCTGTGAACAT	385	

Db	4327	actcttcgcgtcccgccgctcgtatctctctggccaggaacctgccgcggacacttgc	4386
Oy	386	GTCATGATGTACCAACAAGCTTGGTGCCCTCTGGAGSAGAAGATTACGCCCGCATTTTTC	445
Db	4387	atccgcgatccagtcatatccaggacc	4413
Oy	446	ATATTGATCAGCTTCATTGACAGACC	472
RESULT	5		
LOCUS	HUNKFKB34	3205 bp	mRNA
DEFINITION	Human factor KBFL mRNA, complete cds.	PRI	15-OCT-1990
ACCESSION	M55643 M37492		
NID	9189179		
KEYWORDS	factor KBFL.		
SOURCE	Human, CDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea. 1 (bases 1 to 3205)		
AUTHORS	Kleban,M., Blank,Y., Logatz,F., Vandeckerckhove,J.S., Lottspeich,F. Le Bail,O., Urban,W.B., Kouliksky,P., Beaurisle,P.A. and Israel,A.		
TITLE	The DNA binding subunit of NF-kappa-B is identical to factor KBFL and homologous to the rel oncogene product		
JOURNAL	CELL		
MEDLINE	cell 62, 1007-1018 (1990)		
FEATURES	Cell 62, 1007-1018 (1990)		
source	Location/Qualifiers		
	I..3205		

BASE COUNT	874 a	784 c	840 g	707 t	
ORIGIN					
Query Match		2.8%	Score 24:	DB 60:	Length 3205;
Best Local Similarity	78.6%;		Pred. No. 2.03e-01;		
Matches	33;	Conservative	0;	Mismatches	9;
				Indels	0;
				Gaps	0;
Db	1287	ccatttcgagatggttcggcggtgtagtgatgcgcgag	1328		
Cp	168	ccattatctccattctccacggcggtggtggtcccaag	127		
RESULT	6				
LOCUS		HUMNFEB	3625 bp	mRNA	PRI
DEFINITION		Human nuclear factor kappa-B DNA binding subunit (NF-kappa-B)			20-FEB-1991
ACCESSION		M58603			
NID		g189177			
KEYWORDS		nuclear factor kappa-B DNA binding subunit.			

```

ROCHE, BASEL, 4002, SWITZERLAND
FEATURES
    Source
        location/Qualifiers
            1..9152
                /organism="Homo sapiens"
                /clone="L10, L11, L12, and L13"
                /cell_type="B-cell"
                /cell_line="Eppstein-Bar transformed B cell line, also
                called DAN1"
                /sex="Male"
            <1..160
                /note="this is the 3' end of intron 10 that is about 1.7
                kbp long"
            161..299
                /gene="NFKB1"
                /number=11
            join(<161..299,8898...>9041)
                /gene="NFKB1"
                /note="pid:el22791"
                /codon_start=2
                /product="p50/p105 subunit(s) of transcription factor
                NF-kappaB"
                /db_xref="pid:g902633"
                /translation="LPIPSKRSKIKILIONOPICLSSFGNLTWKLNVNLSLTILK
                SKIRKRCGNVRNSCPFRIVSAVVVPELEAECLVVAVEGALEVOVQ"
            300..8897
                /gene="NFKB1"
                /number=11
            360..642
                /gene="NFKB1"
                /rpt_family="Alu"
            8898..9041
                /gene="NFKB1"
                /number=12
            9042..>9152
                /note="this is the 5' end of intron 12 that is about 1.7
                kbp long"
                /number=12
BASE COUNT      2884 a 1682 c 1885 g 2701 t
ORIGIN
    Query Match      2.8%; Score 24; DB 54; Length 9152;
    Best Local Similarity 78.6%; Pred. No. 2,03e-01;
    Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Db 8937 ccaatttcgatatgcttcgcggtgtagtggtgcgcgag 8978
      ||||| ||| ||| ||| ||||| ||||| ||||| |||||
Cp 168 ccaattattccaatcttcacgcgcgtgtgtgtgtgccagag 127
      ||||| ||| ||| ||| ||||| ||||| ||||| |||||

RESULT      8
LOCUS      CELF1BC5      29095 bp      DNA
DEFINITION      Caenorhabditis elegans cosmid F1BC5.
ACCESSION      U28097
NID      9861394
KEYWORDS
SOURCE      Caenorhabditis elegans strain-Bristol N2.
ORGANISM      Caenorhabditis elegans
Eukaryotes: mitochondrial eukaryotes: Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Pseudocoelomata; Nematoda;
Secernentea; Rhabdilita; Rhabdilitida; Rhabdilitina; Rhabditoidea;
Rhabdilitidae; Caenorhabditis.
1 (bases 1 to 29095)
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M.,
Bonfield, J., Burton, J., Connell, M., Cosey, I., Cooper, J.,
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N.,
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,
O'Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roopra, A.,
Saunders, D., Shownkeen, R., Smalton, N., Smith, A., Sonnenhammer, E.,
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,

```

TITLE
 2.2 kb of contiguous nucleotide sequence from chromosome III of C.
 JOURNAL
 Nature 368 (6466), 32-38 (1994)
 MEDLINE
 94150718
 REFERENCE
 2 (bases 1 to 29095)
 AUTHORS
 Fulton, L.
 The sequence of C. elegans cosmid F18C5
 JOURNAL
 Unpublished (1995)
 REFERENCE
 3 (bases 1 to 29095)
 AUTHORS
 Waterston, R.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (13-JUN-1995) Robert Waterston
 COMMENT
 Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63110, USA, and
 Sanger Centre, Hinxton Hall
 Cambridge CB10 1RQ, England
 e-mail: twenemacode.wustl.edu and jess@sanger.ac.uk

NEIGHBORING COSMID INFORMATION:

5' cosmid is C56E6, 200bp overlap; 3' cosmid is C56C10, 200bp
 overlap. Actual start is at bp 1 of this sequence; actual end is at
 bp 2639 of CELC56C10

NOTES:

Coding sequences below are predicted from computer analysis, using
 the program Genefinder (P. Green and L. Hillier, ms in
 preparation).

FEATURES

SOURCE

Location/Qualifiers
 1..29095
 /organism="Caenorhabditis elegans"
 /strain="Bristol N2"
 join(59..115,607..719,867..1020,1064..1214,2717..3300,
 3347..3499,3545..3618,3691..4156,4204..4349,5021..5177,
 5227..6489,6540..6722,6775..6914,6962..7370,7414..7588,
 7573..7880)
 /gene="F18C5.3"
 /map="11"
 /note="coded for by C. elegans cDNA yk29d2.3; coded for by
 C. elegans cDNA yk2a12.3; coded for by C. elegans cDNA
 yk7b7.3; coded for by C. elegans cDNA yk7b7.5; coded for
 by C. elegans cDNA yk2a12.5; coded for by C. elegans cDNA
 yk29d2.5; similar to S. cerevisiae hypothetical 287.5 kd
 protein in PDR region (sp:YBA4_YEAST)"
 /codon_start=1
 /db_xref="pid:g861397"
 /translation="MNADLVKSLRNWVCELAIFHAAPPDROIGKDEMLYIKNVLTNO
 YVEERPLIVIORESEINPIAPFAIVKLLASICRISPEYOMLTVTIMWRKEGKQVLL
 KCDADLLAPLMGCGYDLMEKTRNGILKSLADEPMISTDLVDQCKVORRKR
 SGTSLTLRKRIILEPLIKOIEKVSNNKRPASESTILRESDPVSESDYANLTA
 TPGLICEKRSNEDTLVNMKSLARIPILKDSQYFMRILPFSMBERTLRDLY
 TVEGFKVSMNIDETKTEILYICLIDWMDTKYDEPHFRKRAIIEKILFNGVLA
 TNDHYAMTSSHPHILISTTDLSTRNAGNNVAMQYNGKTLPIAEKILFNGVLA
 PTTLIAFKLDVAVREEMNTLVGATKQFSGSHLAROKRAIKRLVESVAGSKNLSF
 NNNKRLIPMHPILVNFSAFNKISDSGLLILMSKAPWKGKCYLESWLAIREK
 ARAIKDNEFTDALVIVAAVEAHREHVSLEDEKFEISKQNEDESDQSKNKHIL
 OAINRVLPRLSKLSQAOAVERNARAEHTAHATLIDIORAPIALATYKLOKLPD
 AVYGRHLHGVTILKICNLMHTSYSDVETARKTVOIVKCGPYRLASVTEISLWTK
 GEQVAVAFSVTTLVAMREYTKGSELTAINIVKMCIRDOFAADEKEGAVKACCP
 EAKGNAPEMQHLGRIVSPAGIOMVLAFEDVYNEHSPKAAVORVSDCLIKRAGLK
 DNEISDHSQGLDGYIKFSLSDIOKLEVEETNQGDKRGRPRESCILIPAPORIG
 KASKTVISRDSHVAEFVLLFSSLLKEKFDLSDSMVSRILNPFVKIILDCPEFYE
 KISCSMRALSMIOMQFPAJANSORSQDTFLISDYSIGDAGNKRAIIOLNOLI
 YGFTNLIVTTSQDPLDNDKTLILLAYAEADVMDHQOAMFSLIKALVKKVRRHERL
 GIMHLSSETAIRSPLVIRSOCDTLLDYDGSADQAEKVEKVEHVEFIDOLFEYERT
 GNGSAEMLEHLEFKNLVAKSLAPLHMLLVYKMGAAANNESPKVAHIGLALRHLES

CDS

join(8186..8300,8357..8413,8461..9238,9284..9354,
 9408..9786,10106..10191,10498..10580,10631..10847,
 10897..11222,11270..11369,11419..11543,11595..11727,
 11772..11864,11912..12071,12121..12355)
 /gene="F18C5.2"
 /map="11"
 /note="coded for by C. elegans cDNA yk41c3.3; coded for by
 C. elegans cDNA yk41c3.5; coded for by C. elegans cDNA
 yk7h8.3; coded for by C. elegans cDNA yk7h8.5; similar to
 DEAD box helicases"

CDS

complement(join(12967..13103,13609..14014,14063..14260,
 14515..14653,15136..15190,16544..16811,17055..17258,
 17327..17527,17577..17711,17756..17950,18000..18080,
 18064..18362))
 /gene="F18C5.4"
 /map="11"
 /codon_start=1
 /evidence="not_experimental"
 /db_xref="pid:g861398"
 /translation="MATNNSPFRVAKGEVOERKNVNVACDVYNGKITVDVSOLM
 GVLEFADDELINNGETINIKTTADFIRALSAKLPKVIIEFRDHDCLVTEKPMPPRRP
 NTOMVELTFTYRGAAATGIIIRHMAVDVDRKTYTISWOSASVACHYVKGDIIVKANDI
 YTDVRSARKLLAYASVNNKTVRLERSTTDPGAPAPPPDCAQSTPGVAGSGNS
 VMLEPPKPPQGNARKOFEVLLPDDVLEIMKANKNFKYKPCNPCCIIKTAASAPAPV
 PATLPDDPPEVAPVPIPEPKPLKMTPKRSIIIRFPLLVSSQSCSKSKRLRLVY
 LLIISFSFSSCLISCTDETYETAGAAVRYLCAAGCCDPPRAITDTONTEWKKRM
 DELRLIIFPDYGLINNSTIMAPRAAANNEPRDIPADLTDDEDEEOTIPVYN
 OGFAVAVPSRPMOAKKILAVLIVLYFYINDFYKFAVADMODARITPSSIN
 HRDMAVSTYTCOGFDYKKNIMVYKMSGLSTIATLSLITSSPDNKRAMAVI
 RYVDLCAHFIISLEFARISAVTINHTQSLSPASRISVNTLVNDGNSLSTVAPR
 GELPTSRKLIISIFPVQVIEYLITLITGAVIGLAVLIEMCIRHFPPTHBKLTPI
 SAISRLNIVSSSCSSDSNVLA"

tRNA

CDS

complement(join(18753..18876,18926..19369,19411..19477,
 19637..19870))
 /gene="F18C5.5"
 /map="11"
 /codon_start=1
 /evidence="not_experimental"
 /db_xref="pid:g861399"
 /translation="MKLIIAVCYGPAACEVVPASARIPBGDVRRSPLESLSYA
 PPAKTSRPSSTIRAPAVVITLPAQAKRSKRPAAPLPATLPLPLPLPLPLPLPL
 PPAVLSAASDPL
 AMETVATESDNDITSEARASAPYISDNLVITLSRLSKFVRSSEKKVYTAEDNADWI
 VPPH(22223..22486,22538..22683,22771..22846,22947..23316,

CDS

23376..23446,23488..23643)
 /gene="F18C5.1"
 /map="11"
 /codon_start=1
 /evidence=not_experimental
 /db_xref="PID:g861395"
 /translation="MAMYNQTAELTWTQSDGIPKSONSIMKINPVEFLIFL
 TPLSPKARVIRKHNYSKSOILMTITLNLNOLILEIRLRLVHIFINSEDP
 CKIEHSPETIDYITAFISVSTGSLATFDPRFALASTIVYNSKDSAMLT
 VSLIVYIVIRYGVSRAGYVSCYPPQLSNTYQVNNALFWMIMAMCVLTIV
 LNLIVYKDRLLIFEFKLENNQFRIKRSYFDRKTSNFENLITKAICVSSTOF
 VPLSFSTAALAIIRLEAGSEEVFIINICINGYGNLSIPILYIKTNOCILQNR
 KSIDKMTNHTGVDSHISLKTAMET"
 complement(join(23918..24100,24145..24215,24262..24386,
 26110..26528,26759..26986))
 /gene="F18C5.6"
 /map="11"
 /note="similar to F44F4.7; transposon contained within
 intron"
 /codon_start=1
 /evidence=not_experimental
 /db_xref="PID:g861401"
 /translation="MSECCARSDVHNVLJTSNMKFNHCIIISIIISFTTKSVRL
 LKONLPTCTRNLFSAINGIIHOCTAVYRLAFAIYVADPCAILFSSOCPE
 DGNLYVTNLFSSFCFSLFLDRFSRPSRSTVNNHOTLASIVLIIISOILPGLPW
 VFYDATSYVLMCTYPPPMYAKLHEVNNIRICVLYLFAIFLIIHNKIRKRV
 HNYNINSRKSTENYLATKSVCLVITISQICVGPISITSVFIRFDSIPLEWHLI
 ISYLTGLYSNFFLLIILIIODKRIAKRRIMIORLONKNETSPDHDTLSLGMKKT
 GNOETLF"
 complement(24698..25519)
 /gene="F18C5.7"
 /map="11"

Note: remainder of annotations omitted.

Query Match 2.8%; Score 24; DB 26; Length 29095;

Best Local Similarity 92.9%; Pred. No. 2.03e-01;

Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 23819 ttctgtccagagagtgccattt 23846

Oy 508 TTCTGTGCCAGAGAGTGCCATAT 535

RESULT 9
 LOCUS CELB0361 42528 bp DNA INV 26-FEB-1994
 DEFINITION Caenorhabditis elegans cosmid B0361.
 ACCESSION U00031
 NID g458948
 KEYWORDS acid phosphatase; transmembrane;
 phosphatidylserine decarboxylase.
 SOURCE Caenorhabditis elegans strain Bristol N2.
 ORGANISM Eucaryote; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
 Rhabditina; Rhabditoidea; Rhabditidae; Caenorhabditis.
 1 (bases 1 to 42528)
 Wilson, R., Almscough, R., Anderson, K., Baynes, C., Berks, M.,
 Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
 Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,
 Fulton, L., Gardner, A., Green, P., Hawkes, T., Hillier, L., Jier, M.,
 Johnston, L., Jones, M., Kersey, J., Kirsten, J., Laister, N.,
 Latreille, P., Lightning, J., Lloyd, C., McMurtry, A., Mortimore, B.,
 O'Callaghan, M., Parsons, J., Percy, C., Rifken, J., Roopra, A.,
 Saunders, D., Showkeen, R., Smailson, N., Smith, J., Sonhammer, E.,
 Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
 Vaughan, R., Waterston, R., Watson, A., Wellstock, L.,
 Wilkinson-Spoat, J. and Wohldman, P.
 2.2 kb of contiguous nucleotide sequence from chromosome III of C.
 elegans (see comments)
 JOURNAL NATURE 368 (6466), 32-38 (1994)
 MEDLINE 94150718
 REFERENCE 2 (bases 1 to 42528)
 AUTHORS Du, Z.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

The sequence of C. elegans cosmid B0361
 Unpublished (1994)
 3 (bases 1 to 42528)
 Waterston, R.
 Direct Submission
 Submitted (26-FEB-1994) Robert Waterston
 Submitted by:

Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63110, USA, and
 Sanger Centre, Hinxton Hall
 Cambridge CB10 1RO, England
 e-mail: rtw@emacode.wustl.edu and jee@sanger.ac.uk

NOTES:

Coding sequences below are predicted from computer analysis, using
 the program GeneFinder (P. Green and L. Hillier, ms in
 preparation).

FEATURES
 source Location/Qualifiers

1..42528
 /organism="Caenorhabditis elegans"
 /strain="Bristol N2"
 join(670..924,2021..2142,2405..2564,2997..3037,3123..3383,
 3657..3942)
 /gene="B0361.4"
 /map="3"
 /note="coded for by C. elegans cDNA CEESE79FC"
 /codon_start=1
 /db_xref="PID:g458952"
 /translation="MLSLCEFYILFYINICGASHSHHHHSHQHLVNSTLNFSESV
 SNFLSDSGSDLSPLTASNDRRSRSLCSQMKDDQAAVAAQONKDKTGGSGG
 RKASGIVPSLNLRIQCEKFAKPSIGDAIMRAASRAEMTRLSKMEKQIEIG
 KOMEELITDAENADSEKSEYFSGISELITIFWFSHPDSVTHMLYVMAADLIS
 PLIMQOLQNVNMEFIDYKRCIGANGANSILCSLYEVNVIKKEGIRELXRRIL
 SRNSPSTFRKHPTRASDRLKPSITIVSNYSFELKTIISIFVANSFFFEKKRAHF
 FNSSSSFHPGFYVLDKSSYSCYSIIYFYFFSIF"
 complement(join(4177..4486,4678..5207,5737..5861,
 6429..6465))
 /gene="B0361.5"
 /map="3"
 /note="coded for by C. elegans cDNA cm02b7: similar to
 phosphatidylserine decarboxylase"
 /codon_start=1
 /db_xref="PID:g458953"
 /translation="MALNPLIFVAFMSWKMVLSVTLIIIGASVYGVYFTPDREIV
 DSKHYSNKIRIVYLSPEVTASRYVGGLANQDIPYWLREHLGGRARITDCMDCV
 DDPKATPSFAAFNRKLESTRPISASPLVSPADSTVLFHGVEDNKLETVAGHDYD
 VDFLDVDVDPORDELDLYOVVYFLAPGDYHAHSPARVYANQCRVPGILLSVRPTL
 LSHVPHLFCLENERVVLNGSRGFFMSAAVAVNDIVDAPSRITNIVRSKTOY
 MNMTEIHAHYVGERVGERPLSGSTVLVQAPPTIKFAIKADPRLRYGOSLADGV"
 complement(join(11179..11392,11488..11568,11624..11738,
 11844..12220,12301..12463,12515..12604,12653..12749))
 /gene="B0361.6"
 /map="3"
 /note="coded for by C. elegans cDNA CEMS50F: similar to
 H. marismortui hypothetical 23.1 kd protein in HML5 5'
 region"

CDS

CDS

CDS

18311..18691,18769..18849,18904..18997,19045..19313)

```

/gene="B0361.3"
/map="3"
/note="contains transmembrane regions similar to those in
E. coli tetracycline resistance protein and arabinose
transporter."
/codon_start=1
/evvidence-not_experimental
/db_xref="PID:9458951"
/translation="MSISRSRYSDPDMKSENOENNSKKKSSERKLLDPKFEVAYG
AYGQIFIVYVOTLNFYSSSMYINSFVQMLERKOCYKNETIPSTCCIEEES
KAFGNLNGEYCGIENVNTQKASTNLVDFSCSHFEGFICGIVAVIA
VPFMSMLADRGKPIIVTALAFLANMASSPMPAFILIRAFICGSPSYVIA
SVATCEYSEKARAMITVYVNVAMSLGMVTLVLTATDMRMRYVLSLPCYFAL
MYLEPSPHMLITKNTKELKRYIKTANRVIISLYEALSPMSVELGDOVOALYXS
LIEIPAGAVIPLMMKGRKIYIMCIVQTLAIGVYELDSYEFKIVIMVAMVA
TIISVHPVATBOEPTSVSLCFESLMIPOSIGIIMSPYKHVIMSPYKIVYAL
FSFISATLAEVTHETKKKLPDIESTSVSEINDLSAYRKSSSSVSLSTSV
SKTISESVSKIDTNEETEMNSHYKHHTYSTPONCNPFFRRSHVLTORMLI
VNOHPRYTEPLNSTIYFONDCNNINLAMIHNGSAEKSSIDYSDIPRGACQIV
QOLAKNLKLSNMQITFKQWHEVETRIGHEISLSSCQENMSKCTMPVABHEQIV
BUDAKRNLKQLRDLRAQIFDAEDSSEKSLDRAIDDSLPTLVAAGDLKIKTEL
GAKRFDPAETVHKTGRYLAWKQTEFAVFPDPPGLDIDFAPTLNANDEYRLYOH
HFDOSLNNLNKNSGIDEDSKLKMGTQEVYTKAHEFRHIGYSLCKYCDENAFRL
DAMKNYIKNGTLEKLEFQSMENYNYDLSDLKSVYTKAYEEFENTAGDEDEESI
ENIERIIVDGLMESAKKPLSEKLSRMYEPPNFRSVAMSLNFAVILLREISHOL
LEIYLEDGEREFMISILKPIFTDYTRMLCRMEDNSTPKVLEEVLEIALAGOSTIY
EYEDQTEIKNGLELKNKAKMNERVKAIRASHREVTDKMPHNGSEYKESADAKS
ALPTODEVYATONLHLRWEALANENTKVALIAHVMINKSELETPDDAIFA
ALDRISHYVKKLLOSINDVWMLDGTATLSGLVKERLCDAEVLRLADLRAQTH
DNIDSTNKLREOLKTKMS"
complement(join(19585..19730,19811..20069,20113..20325,
20573..20758,20806..20952,21055..21091,21143..21350))
/gene="B0361.1"
/map="3"
/note="coded for by C. elegans cDNA CEBSW24F; similar to
acid phosphatase"
/codon_start=1
/db_xref="PID:9458955"
/translation="MLLLVLLIAGSINAVYKVEPLQANTDLLEVYHVRHGDRT
PAELFDDITKMEGELTQGAQOQLRCLMGLKRRGSLGKPRNRYIRSSD
NRTILMSQANMAGLFPKPIYGLMGMOPIPHITISKPTDELKEEASCTAEEMN
AOMSTKANGIRKKAARELSFESOKLNPMLKATWRFQDLFCENONTTWSMN
SSIFERYDOLYNSOLEPHTDRLRRLRGSTLLEEFHFRSPKASGSKRAKYAS
AHOSTIALLATIGVFYDIPIKYATCLLIMKILANERLIVFEKNETDIDRLIYS
IPGDDPCTLOKIGDDLKRYFEDWAEVVISLIVSYGCTLPEFKRKRLIFR
PVGIDRDTAPMNGDSD"
complement(join(21997..22266,22696..22845,22891..23081,
23364..23470,23685..24031,24082..24353,24615..24725,
24773..24849,25101..25395,25508..25636,25682..25777,
25914..26026))
/gene="B0361.8"
/map="3"
/note="similar to cytoplasmic domain of synaptobrevin"
/codon_start=1
/evvidence-not_experimental
/db_xref="PID:9458956"
/translation="MSDVIISLSHSITVFYVLPILIALIIPESLYSGFRKKTVA
FFHYCNAGGGERVLAIRMOKKEPDKHYFVSGPDMAKEDKILAKORGITL
DPSNIOFIYHMTLIVARHYKCTMLPOLAGLITALEAMRMPALYKIDMSKPSL
LPARLSGKVYVYVHPITSCMDLDVYSROETNENSTTAQSVLWSGLTYRRL
ACLYWAGKRAHYGNSTOHRHTISYRSDYIVYPCVAFELTIESVASLE
DTKYVRLSGQIRPERNHLQLEVDHDKETPELKMGNVBLICGNGRNEDEERYK
MLKWEAKLIDISQILWQNLVPEYDLVLSLITHEEVETILITVHEGKLRNDYKVAASIT
ILNSDGGPRMDIVKIDEGHCVGYLSITLKEEVETILITVHEGKLRNDYKVAASIT
IRGEAAFEVSRGKLKAKNGIFENRKKKULGTGKSKCCENHFOIINIFETPLX
FLVGLFLYNTMLKULSILVFNKNDVSDVLFKSECDLSFSFGFSGVSEPTAL
LVRSRGIGARSYKENEYLVHCVKNDGSLAVCTADAYEOVRVANSKGVRLDFTFR
VPATOWGIRSDKDCSYTGKLDLEKQNPREADPMTROOEVEETKVMYNTIOSVL
DRGKLDLIVKSENSDOSKMEYTSARKMKCNV"
complement(join(28730..28781,29112..29230,29285..29449,29531..29692,
30013..30129,30300..30479))
/gene="B0361.2"
/map="3"

```

*** Note: remainder of annotations omitted.

```

/codon_start=1
/evvidence-not_experimental
/db_xref="PID:9458950"
/translation="MHTFYTHRAIPQIGYGNFRPPDDDDDDSGRYKRYKPKPIAK
YARKLGSEFGKTFVYGSQSKQDFYGSQOVRNDVWSSDSGAFCAVFEHFAVQ
ILKAMKGDIDTLVKRLKRLKLESGIDGDEPKSKRYVTMMRRDRGNTLIPASSRSD
SDRHEGSSRMKREYKESQDLSMREKGTIGDQLRLEFRLSKIRRHDESV
DIIAMOKGKKKSEDEKDRKESIKGVKISEFDEFLKTFEI"
complement(join(34166..34284,34390..34719,34823..34958))
/gene="B0361.9"
/map="3"
/codon_start=1
/evvidence-not_experimental
/db_xref="PID:9458949"
/translation="MTSSPDLHYVYEHKRIERSPDGCSNCRIDSSRLKHNIIAVGN
TYLAVENDGLDDHLIIVPHQSSSTIQIDENVEMMLMRGLVAVWKSQNRICIF

```

```

Query Match      2.7%; Score 23; DB 25; Length 42528;
Matches Local    26; Conservatvie 0; Mismatches 3; Gaps 0;
Db 28059 agaaggaanaactgaataatcgtgata 28087
      ||||| ||||| ||||| ||||| |||||
Qy 199 AGAAGAGAGAACTGAATAATATGTGATA 227

```

RESULT	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
10	HEAH10655	819 bp DNA	BCT	01-OCT-1995	Haemophilus influenzae Rd shikimate 5-dehydrogenase (aroE) gene, complete cds	145396.142023	g1005525	aroE gene.	Haemophilus influenzae (strain Rd) DNA.	Science	269 (5223), 496-512 (1995)	Whole genome random sequencing and assembly of Haemophilus influenzae Rd
1	(bases 1 to 819)	Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.-F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.-I., Hedges, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedlund, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fitchman, J.L., Fuhmann, J.L., Geoghegan, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C.	95350630	Address all correspondence to: J. Craig Venter	The Institute for Genomic Research	932 Clopper Rd, Gaithersburg, MD, 20878	(New address as of Sept 9, 1995: 9712 Medical Center Dr, Rockville, MD, 20850)	Jean-Francois Tomb, Brian A. Dougherty, and Hamilton O. Smith are with the Johns Hopkins University School of Medicine, Baltimore, MD, 21205. Joseph M. Merrick is with the State University of New York, Department of				

JOSEPH M. WEILICKA IS WITH THE STATE UNIVERSITY OF NEW YORK,
DEPARTMENT

of Microbiology, Buffalo, NY, 14214. Chris Fields' current address is The National Center for Genome Resources, Sante Fe, NM, 87505. Keith McKenney is with the National Institute for Standards and Technology, Gaithersburg, MD 20878. All other authors are with The Institute for Genomic Research, 932 Clopper Rd, Gaithersburg, MD, 20878. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332. e-mail: mark@ember.gatech.edu Submitted by: Robert D. Fleischmann e-mail: rdfs@ceitlgr.org Annotation: Owen White e-mail: owwhite@tigr.org Biological role information, putative identifications, sequence alignments, on-line name and sequence search capability are available at TIGR's World Wide Web site (URL: <http://www.tigr.org/>) *** Comment added by NCBI staff: The H. influenzae sequence was re-analyzed with several different methods for protein similarity detection; the resulting set of predicted genes that have homologs in databases was used to re-train the Genmark program in order to predict additional genes among open reading frames without homologs. As a result, some predicted small genes from ref. 1 (with subsequent modifications) whose existence could not be corroborated were discarded, and 23 new genes were predicted. An attempt was made to tentatively reconstruct all coding regions containing frameshifts or in-frame stop codons, including those that have been annotated as misc_features in ref. 1. The sequence was re-annotated based on this analysis. The DNA replication origin (oriC) was chosen as the new zero point. Gene names were assigned to those and only those genes for which apparent orthologs could be identified in E. coli. Address all correspondence concerning the H. influenzae sequence re-analysis and re-annotation to: Eugene V. Koonin or Roman L. Tatusov or National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894 email: tatusov@ncbi.nlm.nih.gov, koonin@ncbi.nlm.nih.gov R. L. Tatusov, A. R. Mushegian, K. E. Rudd, and E. V. Koonin are with the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894 W. S. Hayes and M. Borodovsky are with the School of Biology, Georgia Institute of Technology, Atlanta, GA 30332-0230 P. Bork and N. P. Brown are with the European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900, Heidelberg, Germany.

NCBI gi: 1220729

Location/Qualifiers

1..12500

/organism="Haemophilus influenzae"

/strain="Rd"

/complement(54..1169)

/gene="asd"

/standard_name="HIN0066"

/note="homolog of DHAS_ECOLI, BLAST score: 704; HI0646;

NCBI gi: 1220730"

/codon_start=1

/transl_table=11

/product="aspartate-semialdehyde dehydrogenase"

/db_xref="PID:g1220730"

/translation="WKNGFTGWGNGVSLMDRMSQENDFENLNPFFTSQAQAKA
PVEGKDAIDLKSAFDEIEELKKDIIYTCGGDTYNEVYPKLATGMDYWDASAL
RKDKDAIIVLDVNVQHVISELKGKIDITFGGNGCTVSLMAIGKGFEDLVEMISLA
TYQAAGAGANMRLLSQMGLQEAASSELKPASSIIDIERKTVAKKRRANPPTDN
FGAALGSLIPWIKLPLETQGTKEWKAENKIKGLSDNIPVDGACVIRGLRC
HSOAFTIKKKDLPLEETEQIIASHNEMVAVINDKEITLRELTTPAKVIGTISVPR

CDS

complement(1328..2044)

/gene="yhd"

/standard_name="HIN0067"

/note="homolog of SSU20224_2, BLAST score: 99; HI0647;

NCBI gi: 1220731"

/codon_start=1

/transl_table=11

/product="Mg2+-transporting ATPase"

/db_xref="PID:g1220731"

/translation="MENSILLTALFNPDLIFSKMLLAVIGSGVIGPELKKRPV
GVKCAIATVTCVLTIVSIOAEHVAQSPMPRIAAQVYSGTIGTGACVITH
KNDNISGLTAAITWASGIGIAGAGVPAVATATWIVTSIRLSPLVQWYHRS
ORRRKNTLINDAESIGKVTOLVNNQIRIHIVQKDSGSEVALQIRCSIDSTWL
KDAVALAKAPDCYSVEPN"

complement(2322..2948)

/gene="mdaB"

/standard_name="HIN0068"

/note="homolog of MDAE_ECOLI, BLAST score: 633; HI0648;

NCBI gi: 1220732"

/codon_start=1

/transl_table=11

/product="modulator of drug activity"

/db_xref="PID:g1220732"

/translation="MATEPFIIICFRSKNNIILLDGKAFSGHSELNHTLAKKAK
EVLTAAGHNVKREVIDAGVDVABELEKLMMDAVYQWPMSEHPEPTVKKYDEVLT
NGHGLYHSGRHSVPTEGVTGGLGCKHMTSTWNAETAEPTREGCFPEGKVD
VLVHHRKINERFGLRLPTEFCNVVSPQEVQIADYQAHLEKVEG"

complement(3190..5202)

/gene="rep"

/standard_name="HIN0069"

/note="homolog of REP_ECOLI, BLAST score: 1657; HI0649;

NCBI gi: 1220733"

/codon_start=1

/transl_table=11

/product="DNA helicase"

/db_xref="PID:g1220733"

/translation="MKLNPOQQAWEVYTGCLVLVLAGSGKRTVIINKIAHLIEKCG
YSPKQIAVTFNKAERKERVASHISEQSGILVSTPHTGFDILKREYALGFK
SNMTLDEHDOFALKELTADVLKEDKDLBELLSVISMKNDSLSPKQAFALARK
YQFACQYRYVATRTYNALDDDLILPLTLKQNEDEVRKQAKIRIYLDEVQD
TNSOYELIKLVGRACQTVGDDOSIVSMBRANENVRNRPDPRILNKLQDN
YRSTORILICANILINDNEHVPDKLTSTICKGKILYAKKNEHAETVABELAH
RESRKRYDVAIILYKGNHOSLEKULYQMRIRYKISGGSFSSRELIQOMAYTLN
VYNODDAPFIVNTPKREISTATLQGLGAEKHSLEAFEEELIQTITPAY
DSIQKGRVIEINDIOSEPERVAVRMSLAIIEEYLVETATSPRAAMOSKNVAT
LRFVADMKRGDEINSPMLNNOVATRLRLDMLRGEDDDSDOVQMLTAAKSGLEF
PYIYLIGMEGILPHQTSIDEDNVEERRLVAVITRAOKELTFESLCRERROYGELVR
PEPSRPLAELPNDVYLWERDKPLTTEOKOKETONOLRLAILKS"

complement(5208..5420)

/standard_name="HIN0070"

/note="weak similarity to YEJM_ECOLI, BLAST score: 60;

HI0650; NCBI gi: 1220734"

/codon_start=1

/transl_table=11

/product="H. influenzae predicted coding region HIN0070"

/db_xref="PID:g1220734"

/translation="MKIIFITLALIVSICGGSVYKILDPTEKTYAGVADLEMAQ
OMGLPTLIDLSFLIDVILVYANAO"

complement(5417..5887)

/gene="kdsB"

/standard_name="HIN0071"

/note="homolog of KDTB_ECOLI, BLAST score: 441; HI0651;

NCBI gi: 1220735"

/codon_start=1

/transl_table=11

/product="lipopolysaccharide core biosynthesis protein
(cytidilyltransferase ?)"

/db_xref="PID:g1220735"

/translation="MTSVIYQKEDPPLTNGHLDIIEASVIFPRVAVANSPSKPL
FSLSEVEVLRVSVHLSNVEPSPDLNANVYKQNHISLITNGVTTTIDEFEIOLA
ANRLITGVDSFPFPAKFAFVSSTIVREITLHGQDAELVAVPFFNALKAR"

complement(5884..7167)

CDS


```

FT      /db_xref="pid:91220731"
FT      /translation="MENSLLLTALFNPDLIIIFSKMLLVAGVIGLELELRKRPV
G      VETCALIANTTCVLTIVISIQAEHQAQVSEIRTPDKRLAQAQVIGIGFAGVILHK
FT      NDAISULTAIIIMASAGIGIAGAGVFDVAIVTMIIVSIRLSPVORWHRKSQR
FT      KRFENILVDAESIGKVTQLLVNQYRIEHIQVQDSSGEVRLQIRCFSIDSTMLKDA
FT      ALLKADGVISYEVDN"
FT      complement(2322..2948)
FT      /gene="mdab"
FT      /standard_name="HIN0068"
FT      /note="homolog of MDAB_ECOLI, BLAST score: 633; HI0648;
FT      NCBI gi: 1220732"
FT      /codon_start=1
FT      /transl_table=11
FT      /product="modulator of drug activity"
FT      /db_xref="pid:91220732"
FT      /translation="MATPFFIILICFSKKNMILLDGGKAFGSHGELNHLHKAK
E      VLTALGHNVKEIVDAGYDVAEIEKFLMDAVIMQNP SWMHHEPMTVKYIDEVLTN
G      HGKLYHSDGRHSVNPTEGYGTGLLGKRRHMLSTTNAPLEAFTREGDFEFGKGYDL
FT      MHFRLNFEGLTLPFLCNDVYKSPQVEQYLADYQAHLEKVEG"
FT      complement(3190..5202)
FT      /gene="rep"
FT      /standard_name="HIN0069"
FT      /note="homolog of REP_ECOLI, BLAST score: 1657; HI0649;
FT      NCBI gi: 1220733"
FT      /codon_start=1
FT      /transl_table=11
FT      /product="DNA helicase"
FT      /db_xref="pid:91220733"
FT      /translation="WKLNPQQQQAWEYVTPGCVLACAGSGKRVYINKIMHLIEKCG
Y      SPKQIAAVTFTNKAREKERVANHSICKESKGLVSTFHTLGFDLIKREYALGPKS
N      MTLFEDHDOFALLKELTADVLKEDKDLLRELISVSNKNDLISPKQAFALARDAKYO
T      FAKCYERYATOIRTYNALDFDDLIMLPTLLFKONEEYRSKQAKIRYLLVDEYODTNT
S      QYELIKILVGRACFTVVGDDDSIYSMRGAPENNVRLRDPRLNVIKILBQNVRS
FT      RLHCANILIDNNEHVPDKLFTIGKEKLVITAKNBEHEAERIVALLIHRFSRK
Q      KYKQVAILIRGNHOSRLEKVLQNRIPYISGTSFSSAEIKDKMAVLRVNVQDD
T      AAFLRIVTPPREIGTATLQKLGELAQEKHISFEALFEPELQRTTPKAYDSLQKFG
D      AAFRLIIVTPPREIGTATLQKLGELAQEKHISFEALFEPELQRTTPKAYDSLQKFG
R      WIVELNDEIQSEPEPRAVRSMLSAIIHEEYLYEATSPRAEMQSKNVATLFDWVADM
L      KQDETNEPMINQVYVTRLLTLDMLERGEDDDSDQVQMLTLLHASKGLEPYYVLIGME
L      GILPHQTSIDEDNVEERRLAYVIGITPAQKELTFSLCRRROYGELVRPEPSFAEL
E      NDVLMERDKPKLTTEOKOERTQNOQLRLRAILKS"
FT      complement(5208..5420)
FT      /standard_name="HIN0070"
FT      /note="weak similarity to YEJW_ECOLI, BLAST score: 60;
FT      HI0650; NCBI gi: 1220734"
FT      /codon_start=1
FT      /transl_table=11
FT      /product="H influenzae predicted coding region HIN0070"
FT      /db_xref="pid:91220734"
FT      /translation="MIIKIFLTAIIVLSGCGSVKLDPTKRYTAVAGVADLEMAQ
P      WGLPILDLPLSLDITVILPYAMAQ"
FT      complement(5417..5887)
FT      /gene="kdtb"
FT      /standard_name="HIN0071"
FT      /note="homolog of KDTB_ECOLI, BLAST score: 441; HI0651;

```

```

FT      NCBI gi: 1220735"
FT      /codon_start=1
Note: remainder of annotations omitted.
Query Match      2.68; Score 22; DB 10; Length 12500;
Best Local Similarity 73.98; Pred. No. 4.06e+00;
Matches 34; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
Db      8749 tttaaatccagcatcgaagccgttcacacactaatcga 8794
Cp      202 TTCATAAATTCTGATCATCCAGCGCTGTCCACCAATTATTCGA 157
RESULT 14
LOCUS      HUI32694 12500 bp DNA 29-FEB-1996
DEFINITION      Haemophilus influenzae aroe, asd, kdtA, kdtB, mdab, rep, tag, yhes,
yhd, yrdC, yrd genes from bases 87095 to 99594 (section 9 of 163)
of the complete genome.
U32694 142023
ACCESSION      91220729
KEYWORDS      complete genome; shotgun sequencing.
SOURCE      Haemophilus influenzae.
ORGANISM      Haemophilus influenzae
Eubacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
1 (bases 1 to 12500)
Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A.,
Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.-F.,
Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., Fitzhugh, W.,
Field, C.A., Gocayne, J.D., Scott, J.D., Shiley, R., Liu, L.-I.,
Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T.,
Heublen, D.M., Brandon, R.C., FINE, L.D., Fritchman, J.L.,
Fuhlman, J.L., Geoghegan, N.S.M., Gnehm, C.L., McDonald, L.A.,
Smith, R.V., Fraser, C.M., Smith, H.O. and Venter, J.C.
Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd
Unpublished (1995) see COMMENT for author addresses
2 (bases 1 to 12500)
White, O.
Direct Submission
Submitted (27-JUL-1995) to GSDS by: Owen White
3 (bases 1 to 12500)
Tatusov, R. L., Moshegian, A. R., Bork, P., Brown, N. P., Hayes, W. S.,
Borodovsky, M., Rudd, K. E. and Koonin, E. V.
Metabolism and evolution of Haemophilus influenzae deduced from a
whole genome comparison with Escherichia coli
Curr Biol. in press (1996) In press
Address all correspondence to:
J. Craig Venter
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD, 20878
(New address as of Sept 9, 1995: 9712 Medical Center Dr, Rockville,
MD, 20850)
Jean-Francois Tomb, Brian A. Dougherty, and
Hamilton O. Smith are with
the Johns Hopkins University School of Medicine, Baltimore, MD,
21205.
Joseph M. Merrick is with the State University of New York,
Department of Microbiology, Buffalo, NY, 14214. Chris Fields' current address
is
The National Center for Genome Resources, Sante Fe, NM, 87505.
Keith McKenney is with the National Institute for Standards and
Technology, Gaithersburg, MD, 20878.
All other authors are with The Institute for Genomic Research, 932
Clopper Rd, Gaithersburg, MD, 20878
Predicted open reading frames were determined using Genemark
software, kindly supplied by Mark Borodovsky,
Georgia Institute of Technology,
Atlanta, GA, 30332. e-mail: mark@amber.gatech.edu
Submitted by: Robert D. Fleischmann

```

e-mail: rdfs@tigr.org

Annotation: Owen White

e-mail: owhite@tigr.org

Biological role information, putative identifications, sequence alignments, on-line name and sequence search capability are available at TIGR's World Wide Web site (URL: <http://www.tigr.org/>)

*** Comment added by NCBI staff:

The H. influenzae sequence was re-analyzed with several different methods for protein similarity detection; the resulting set of predicted genes that have homologs in databases was used to re-train the Genmark program in order to predict additional genes among open reading frames without homologs. As a result, some predicted small genes from ref. 1 (with subsequent modifications) whose existence could not be corroborated were discarded, and 23 new genes were predicted. An attempt was made to tentatively reconstruct all coding regions containing frameshifts or in-frame stop codons, including those that have been annotated as misc features in ref. 1. The sequence was re-annotated based on this analysis. The DNA replication origin (oriC) was chosen as the new zero point. Gene names were assigned to those and only those genes for which apparent orthologs could be identified in E. coli. Address all correspondence concerning the H. influenzae sequence re-analysis and re-annotation to:

Roman L. Tatusov or

Eugene V. Koonin

National Center for Biotechnology Information,

National Library of Medicine, National Institutes of Health,

Bethesda, MD 20894

tatusov@cbl.nlm.nih.gov, koonin@cbl.nlm.nih.gov R. L.

Tatusov, A. R. Mushegian, K. E. Rudd, and E. V. Koonin are with the

National Center for Biotechnology Information,

National Library of Medicine,

Bethesda, MD 20894

W. S. Hayes and M. Borodovsky are with the

School of Biology, Georgia Institute of Technology, Atlanta, GA

30332-0230

P. Bork and N. P. Brown are with the

European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900,

Heidelberg, Germany.

Location/Qualifiers

1..12500

/organism="Haemophilus influenzae"

/strain="Rg"

/complement(54..1169)

/gene="asd"

/standard_name="HIN0066"

/note="homolog of DHAS_ECOLI, BLAST score: 704; HI0646"

/codon_start=1

/transl_table=11

/product="aspartate-semialdehyde dehydrogenase"

/db_xref="PID:g1220730"

/translation="MKNVGIGRMGMSVGLDMRMSQDENLNPVEFTTSQAGOKA

PVEGSDAGDLSAFIDELKRLDIIVCGDGYDNEYVPRKATGMGYWADASAL

RMKDIAIIVDPVNOHVTSEGIKGIKTPGNCNGVSLIMAGISLEPKDIDVEWISVA

TYQASGSGAKMRRELISQKGLLEDAVSSELDPRASSTLIDIERKATARKADNFTDN

FGAAGGSLIPWIDKLPEPTGQTKEMVGAETNKILGLSDNPVPSGLCVRIAGLRG

HSQAFTRIKRDLPLEIEIOIASHNEMVKVTPNDKEITRLPARKVTGTLSPVGR

LRLKAMGPYLAFTVGDQLLGAEPVRRILKQVYA"

/complement(1328..2044)

/gene="yhd"

/standard_name="HIN0067"

/note="homolog of SSU20224_2, BLAST score: 99; HI0647"

/codon_start=1

/transl_table=11

/product="Mg2+-transporting ATPase"

/db_xref="PID:g1220731"

/translation="MENSLLITLAFNDHLIFSKMLAWLGSVIGLERELKRPV

GVKTAIAVTCTVITVSIQAERYAVSENRIDPRLAAQVIVSIGFAGAVILH

KRNPASGLITVAIIWASGIGIAGAGFVDVATATWIVSIRLSPVORWHRKS

ORRTFENILVNDASIGKVTGLVNNQVRIEHIQVDDSGSEVALQIRFSIDSTML

KDAYALLKAEQVIVSEVDN"

/complement(2322..2948)

/gene="mdab"

/standard_name="HIN0068"

/note="homolog of MDAB_ECOLI, BLAST score: 633; HI0648"

/codon_start=1

/transl_table=11

/product="modulator of drug activity"

/db_xref="PID:g1220732"

/translation="MATPFVITICPSKKNMILLDGGAFSGHSHGLNHTLKKAK

EVLALAHNWKETVIDAGYVEAIEERFLMDAVIWMGPMWHEPVTYKVIDEVL

NGHGRKLHSGDRHVNPTGEGTGLGKRRKHMVSLTWNPATIEFTREGDFEFGVD

VLVNHFKLNEFLGTRLPFLICDVVKSQVEGLADYDVAHLEKVEG"

/complement(3190..5202)

/gene="rep"

/standard_name="HIN0069"

/note="homolog of REP_ECOLI, BLAST score: 1657; HI0649"

/codon_start=1

/transl_table=11

/product="DNA helicase"

/db_xref="PID:g1220733"

/translation="MKLPQQAQVEYVTCPCVLVLAGSGKRTVIINKIAHLKCG

YSPQIAVFTNKAAEMKERVASHSGKQSLVSTHTLGLDLKREYKALGK

SNMFLPEHDQFALKELADVLEKEDDLRELIVSNKNDLISPKQAFALARDK

YQPAKCYERATQIRTYNALDFDLJMLPILKOMEVSKQAKIRYLLVDEYOD

TNTSOYELIKLVGDRACFTVYGGDOOSISWBGARBNVRLDDPPRLVLEON

YRSQRLHCANLIDNNEVPEPKKSTIGKELVIAKANEHEAERTVALLIAH

RFSKRTIKDYAILIRGNHOSRLKELMQRIRPKYISGGSFSPRAEDKMAVLR

VVNDDDAALFRLVNTKRELGTATLKLAEKRIISLEALFEELQRIKAY

DLQKFGWVLELDETORSEPERAVSMLSAHEYLEYATSPAAEOKSNVAT

LFEDWALMKGDEENPMNLNOVYRTLRLDMLERGGDDSDOVQMTLHASKGLEF

PYVILKMEGELPHQTSIDEDNVEERRLAYVITPAQKELTSLICRRRQYELVR

PEPERFLAELPNDVLMERDKRKLITQKQEKTNQNDRLRAILIKS"

/complement(5208..5420)

/standard_name="HIN0070"

/note="weak similarity to YEJM_ECOLI, BLAST score: 60;

HI0650"

/codon_start=1

/transl_table=11

/product="H. influenzae predicted coding region HIN0070"

/db_xref="PID:g1220734"

/translation="MKIFIFLALIVLVSQGSVVKLIDPTEKATAYAGVADLEMAQ

QMGFLIDLPSPFLDVLPLPYMAQ"

/complement(5417..5887)

/gene="kdb"

/standard_name="HIN0071"

/note="homolog of KDB_ECOLI, BLAST score: 441; HI0651"

/codon_start=1

/transl_table=11

/product="11polysaccharide core biosynthesis protein

(cytidyllyltransferase 2)"

/db_xref="PID:g1220735"

/translation="MTSVIYPGTDPITNGHLDIERSAVIFPRVLVAVANSPSKRL

FSLEERVELRQSVVHLNVEFGSLLANVIOHISAIIRVRTTTFEYELQLA

ALNLRILKGVSLFFPPAEKWAFFSSTVIREIYVHGDAVELVPPVFNALKAR"

/complement(5884..7167)

/gene="kda"

/standard_name="HIN0072"

/note="homolog of KDA_ECOLI, BLAST score: 914; HI0652"

/codon_start=1

/transl_table=11

/product="3'-deoxy-d-manno-octulosonic-acid transferase"

/db_xref="PID:g1220736"

/translation="MMRFYISLLIOLPLICITGLISVKSPPYRORLARERFYGN

ASCPPOGIFTHASVGEVIAATPLVQLQDDYPHLSITFTFPTPSSEKVAFGGS

VFHYLPDLPEFSIHRFVNEVQPKIVMEELPNLIHQITFPIPEFVARNARLSR

SAHRYGRIKALQDMSQISLIAQDNISGRVYATLGAPEKRLNITGNIKDKLWNE

LKRTIDSRLTLKQDPRIMIAASTHNEDEILKSHALLAKYNNILLILVPRPERP

NVVADLKKKEFGIIRSTNLEPNENOVITIGDSGELMLMYGSDIAFVGSVYKRG

CHNDEPLAEKMPVITIGKHTFPEFIRMLAEVGVLEVNSTADLERAVALLNSKE

SRERLGNAGYVLENRGALQRLDLKPYLERNV"

Note: remainder of annotations omitted.

Query Match

2.6%; Score 22; DB 20; Length 12500;

FEATURES

source

CDS

CDS

CDS


```

/codon_start-1
/product="R06F6.5"
/db_xref="PID:9577765"
/translation="MESHNLQNTSGRNSNDLNNSTISNFTPEVQSTPALLFGKRA
TVPSSYTPASPLNTASAPCSDIFAVASAPVPOHLMDTPGSKSVHNSPSLVOGSKSAO
TONTNPALISFGNNSFAPTKPAPOSISOTPSFGQAMAPPLRLRDKEPAKISRR
NTFARSTPLSTPITORVTSRLAEQPEEADADTWVTFGFOPOSOVSTILNF
SRGEVSSHOTPSKGNFIHWRSCVTAQQAISRNGCLLDQPTFGVOCSTKNDYNG
SASGIVARRSSNIAANRNASMNSFTVENDADQSVNHNENSVLNSNVDFDANNSLNS
SRISVSGVGRRPLAADQRTNGTPSVKRAKPDGLNKFWNTIGLN"
join(20156..20257,20501..20589,22181..22430,22479..22615,
22668..22989,23996..24125,24266..24351,24417..24551)
/note="homeobox protein, which possibly is the homolog of
the drosophila rough gene (Swiss Prot accession number
P10181)"

```

```

/codon_start-1
/product="R06F6.6"
/db_xref="PID:9577766"
/translation="MHRKGOTRLTYHFLHLSPHHSPFLDSIWORMVHPSAFHP
YTRPTTSGHDLPEVDSLRLLLSAEALIALAQADASKITPSTPYKDRSPOLPAM
TPPIATPSTPEQPPQSPAPNERSRRKRTFSPQATRLAEYIGDSYMAKERK
HLAQSLKLSNQYTWONRRAKDRKSEKSNASNHTSNRSPSRKSSSDSTPP
TOAQFDMPTQIQASPTTADSAIFPTSPESIIOKIEQPSNQILPNFDILQTYLQ
SLSSQIPLQFVSTPPLFDPMQLHYIFEFSQSECSYFQNSASNNRSLILY
FFATIFAVDAQYDSDSSQEDLPDIAPESYSEIIEVDVNIQSGSGDGPILDOYH
MFCILPCPGDIRKRGILIEIGRK"
join(25838..25889,26496..26611,26664..26858)
/codon_start-1
/product="R06F6.7"
/db_xref="PID:9577767"
/translation="MRSMIPLVLFAVIAVFAQKSSQSEDSRRRPSKSSSSSDS
KSSDSSSSSESGDYSEAPNTDSTPEILAAKPDSCILLGPDNRVAKRDLPSF
YDKRKRLGPEAYDIRK"

```

```

complement(join(28010..28079,28137..29078,29714..29846,
29895..30012,30062..30281,30682..31664,32263..32554,
32651..32943,33071..33608,33689..33780,33835..34296,
34347..34457,34507..34647,35137..35346,37261..37597,
37771..37877,37930..38018,38064..38506,38552..38729,
38780..38841))
/note="CDNA EST CEEST88F comes from this gene; CDNA EST
yk109.3 comes from this gene; CDNA EST yk21h11.3 comes
from this gene; CDNA EST yk109.5 comes from this gene;
CDNA EST yk21h11.5 comes from this gene; CDNA EST yk47c1.5
comes from this gene; CDNA EST yk47c1.3 comes from this
gene; CDNA EST yk88h2.5 comes from this gene; CDNA EST
yk92d6.5 comes from this gene; CDNA EST yk92d6.3 comes
from this gene"

```

```

/codon_start-1
/product="R06F6.8"
/db_xref="PID:9577768"
/translation="MDESEKNPASPTTGREEELPGSSPEGYPADEHIFHEDQAPLR
VESAKHEEIVEOQOQPEDLEQGDIVEDQOQFMNVQITQEDMEAGFDVAGFE
LNLLEQLNIVVAISOORAKONEQNEVEVLEDGSHHMHHEMENOFEEDNG
PEEYDNGGOIIDNANHILINDGVNTSKOFOYVSPSEIANLINDIANLSTENVH
QVOLALPRIKEKQDSAYNDAPSTSYHHHHHQLAGKSTRPIIGETVQIQTADG
RLQDAVVKYRGDSEIKQLNGSEFATITDQMLVPODRSDHEIQVAPVLRKTD
MASVNAQAQKRASANDDLCPVLKRSYQALPVDPGPHLVHTPNFCPCIDKKVYKEPS
YIVIRLPACDSCITHGRSPRVSVCKNYKGTIGILVYVILKTGOONSRSRENTPK
TGORDKRDTGIDTLVHGNYKMHNTLSTNDPSYLOPKHEKSDTAADIKSYVANDRR
LIAVATNDAYIIMLNPOLLICSVGIDANKETRGELKEIYKMPDSDSTIAVTNOC
ILIYNLDLRDEQCYNTDSADPYFQNRSPELFTKGSRTAHLHPTIINLADIPTC
VPSRDEFLVCQNGFTHVNTGETIASLSLRASSIPSPVQLOSKSENTITSKSTYIF
DAVYAPLGGFAIVLSDQALITSNDPNFAPNALIGVWAPNMKMDATCCVNHKFLLI

```

Note: remainder of annotations omitted.


Query Match 2.68; Score 22; DB 26; Length 41452;

Best Local Similarity 92.38; Pred. No. 4.06e+00;

Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 11613 aataaacagaaaatggcgatatt 11638
|||||
Cp 960 AATAAACAGAAAACGGCCATATT 935

Search completed: Mon Feb 10 23:54:38 1997
Job time : 637 secs.


 王 衛
 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K
Distribution rights by IntelliGenetics, Inc.

```

Mpsrch_mn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on:      Mon Feb 10 23:54:57 1997;      MasPar time 88.81 Seconds
Tabular output not generated.      833.805 Million cell updates/sec

```

Title: >US-08-674-311-1
Description: (122-970) US08674311.seq (2 of 2)
Perfect Score: 849
N.A. Sequence: 122 ATGGCCTCTGGCACCACCCAC.....CTGTTTATTCACAGCAT 970
Comp: TACCGAGACCCTGGGGGTG.....GACCAATAATAGTTCTGTA

Scoring table: TABLE default
Gap 6

```
Nmatch   STD :   Dbase 0; Query 0
Searched: 113505 segs, 43611913 bases x 2
```

```
Post-processing: Minimum Match 0%
                  Listing first 45 summaries
```

```
Database: n-geneseq25
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22
```

Statistics: Mean 8.967; Variance 4.877; scale 1.839

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description	Pred. No.
1	240	28.3	2763	19	T15167	Methylthiodenosine-p	9.50e-15
2	240	28.3	2763	15	O93813	Human <i>MTase</i>	9.50e-15
3	240	28.3	2784	16	O99202	<i>Pseudomonas putida</i> me	9.50e-15
4	73	8.6	1047	2	O10572	Human <i>Nutritive</i> pep	2.34e-13
5	64	7.5	1047	2	O10572	Human <i>Nutritive</i> pep	4.78e-12
6	44	5.2	91	9	O51746	Oligonucleotide probe	1.45e-13
7	42	4.9	91	9	O51746	Oligonucleotide probe	2.82e-12
8	40	4.7	204	1	N81164	Base substituted <i>E.co</i>	5.29e-11
9	37	3.7	204	1	N81164	Base substituted <i>E.co</i>	1.02e-09
10	31	3.4	114	12	O70469	Generic DNA sequence	1.70e-05
11	30	3.5	114	12	O70469	Generic DNA sequence	6.50e-05
12	30	3.5	114	12	O70465	Generic DNA sequence	6.50e-05
13	30	3.5	114	12	O70468	Generic DNA sequence	6.50e-05
14	30	3.5	114	12	O70467	Generic DNA sequence	6.50e-05
15	30	3.5	114	12	O70470	Generic DNA sequence	6.50e-05
16	30	3.5	114	12	O70467	Generic DNA sequence	6.50e-05
17	30	3.5	114	12	O70465	Generic DNA sequence	6.50e-05

C	18	28	3.3	114.12	070466	Generic DNA sequence	9.06e-04
C	19	28	3.3	114.12	070468	Generic DNA sequence	9.06e-04
C	20	28	3.3	114.12	070466	Generic DNA sequence	9.06e-04
C	21	28	3.3	114.12	070470	Generic DNA sequence	9.06e-04
C	22	27	3.2	66.21	113585	TSAR-9 library genera	3.29e-03
C	23	27	3.2	69.21	113583	TSAR-9 library genera	3.29e-03
C	24	25	2.9	74.21	113613	DC43 TSAR library gen	4.09e-02
C	25	25	2.9	75.21	113612	DC43 TSAR library gen	4.09e-02
C	26	25	2.9	81.21	113611	DC43 TSAR library gen	4.09e-02
C	27	25	2.9	82.21	113610	DC43 TSAR library gen	4.09e-02
C	28	25	2.9	114.12	070473	Generic DNA sequence	4.09e-02
C	29	25	2.9	114.12	070472	Generic DNA sequence	4.09e-02
C	30	24	2.8	82.21	113610	DC43 TSAR library gen	1.39e-01
C	31	24	2.8	114.12	070472	Generic DNA sequence	1.39e-01
C	32	24	2.8	114.12	070471	Generic DNA sequence	1.39e-01
C	33	24	2.8	114.12	070471	Generic DNA sequence	1.39e-01
C	34	24	2.8	114.12	070473	Generic DNA sequence	1.39e-01
C	35	23	2.7	74.21	113613	DC43 TSAR library gen	4.63e-01
C	36	23	2.7	75.21	113612	DC43 TSAR library gen	4.63e-01
C	37	23	2.7	81.21	113611	DC43 TSAR library gen	4.63e-01
C	38	23	2.7	565.6	035072	HCV envelope region n	4.63e-01
C	39	22	2.6	65.21	113586	TSAR-9 library genera	1.50e+00
C	40	22	2.6	68.21	113584	TSAR-9 library genera	1.50e+00
C	41	22	2.6	69.21	113583	TSAR-9 library genera	1.50e+00
C	42	22	2.6	498.3	N50034	Sequence encoding new	1.50e+00
C	43	22	2.6	501.3	N50030	Sequence encoding new	1.50e+00
C	44	22	2.6	565.6	035072	HCV envelope region n	1.50e+00
C	45	22	2.6	657.7	043519	Degenerate FMN reduct	1.50e+00

ALIGNMENTS

RESULT	1
ID	T15167 standard; DNA; 2763 BP.
AC	T15167;
DT	29-JUN-1996 (first entry)
DE	Methylthiodenosine-phosphorylase gene.
KW	Human; methylthiodenosine-phosphorylase; chromosome walking;
KW	Interferon-alpha; CDK4; tumour suppressor; chromosome-9p21;
KW	Cyclin-dependent kinase-4-inhibitor; probe; mutation;
KW	Leukemia; melanoma; diagnostic; dysplastic nevus syndrome; gliom
KW	non-small cell lung carcinoma; cancer; gene therapy; antisense;
KW	ribozyme; antibody; imaging; ss.
OS	Homo sapiens.
PH	Key
FT	Location/Qualifiers
FT	254..421
FT	/*tag- b
FT	422..615
FT	Intron
FT	/*tag- c
FT	616..720
FT	exon
FT	/*tag- d
FT	721..963
FT	Intron
FT	/*tag- e
FT	964..1203
FT	exon
FT	/*tag- f
FT	1204..1203
FN	M09528169-A1.
PR	26-OCT-1995.
PR	12-APR-1995; 104655.
PR	14-APR-1994; 05-227800.
PA	(REGC) UNIV CALIFORNIA.
PA	Carson DA, Nobori T;
PI	WPI; 95-373630/48.
DR	Cyclin dependent kinase inhibitor gene, related predisposition and
PT	antibodies - useful for diagnosis, assessing predisposition and
PT	treatment of cancers
PS	Example 1; Page 96-101; 129pp; English.
CC	The sequence encodes a methylthiodenosine-phosphorylase, and is
CC	located at chromosome-9p21. A cyclin-dependent
CC	protein-kinase-4-inhibitor (CDK4) tumour suppressor gene
CC	(115157-58) is located between this gene and an interferon-alpha
CC	gene cluster, and has been isolated by chromosome walking. The
CC	CDK4i gene, probe and primer derivatives and the gene product may
CC	be used in diagnosis of cancer, particularly melanoma (especially

CC dysplastic nevus syndrome), glioma, non-small cell lung carcinoma
 CC or leukemia. The gene may also be used in cancer gene therapy, or
 CC in antitumor antisense oligonucleotide or ribozyme construction.
 CC Antibodies against CDK4 may be used in diagnosis or in vivo imaging.
 SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T;

Query Match 28.3%; Score 240; DB 19; Length 2763;
 Best Local Similarity 99.2%; Pred. No. 9,506-159;
 Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 aggtcttataagactgtctaaagactagactcggctgcctcctcaaggagcaatgg 1021
 |||||
 QY 570 AGGTCTTATAGAGACTGCTAAGAGCTGAGCTCGGTCCACTCAAGAGGACATGG 629
 |||||
 Db 1022 tcacaatcgaggagactcgttttaagctccggcgagaaagcttaagctcgcacctggg 1081
 |||||
 QY 630 TCACAAATCGAGGAGCTCTTTAGCTCCGGCGAAGAGCTTATGTTCCGACCTGGG 689
 |||||
 Db 1082 gggcgagatgtatcaacaatgacacagttccagagtggtcttctgtaaggagctggaa 1141
 |||||
 QY 690 GGGCGGATGTTATCAACATGACCACTCCAGAGTGTTCTTGTCTAAGAGGCTGGAA 749
 |||||
 Db 1142 ttgttacgcaagtatcgacatcgagcagagattatgactgtggaaggagcagaggaag 1201
 |||||
 QY 750 TTGTTCACGCAAGTATCGCATGCGGACAGATTATGACTGCTGGAAGAGACAGAGGAG 809
 |||||
 Db 1202 cagt 1205
 |||||
 QY 810 CAGT 813

RESULT 2
 ID 092813 standard; DNA; 2763 BP.

AC 092813;
 DE 05-DEC-1995 (first entry)
 DT Human MTase.
 KW MTase: methyladenosine-phosphatase; malignancy: ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 254..421
 FT /tag- a 616..720
 FT exon b 964..1203
 FT /tag- c 1640..1762
 FT exon d 2272..2310
 FT /tag- e
 FN M09318233-A.

PD 06-JUL-1995.
 PR 22-DEC-1994; 014920.
 PR 29-DEC-1993; US-176855.
 PA (CIBA) CIBA GEIGY CORP.
 PA (REGC) UNIV CALIFORNIA.
 PI Carson DA, Nobori T, Takabayashi K;
 WP1: 95-246398/32.
 PT Detecting methyladenosine phosphatase in mammalian cells - by
 PT hybridisation with specific oligonucleotide for detecting malignancy,
 PT also new nucleic acid, expression vectors, derived polypeptide(s) and
 PT antibodies
 PS Claim 8: Page 34-35; 47pp; English.
 CC A cosmid gene library constructed from human placenta DNA was
 CC screened using a MTase cDNA probe to isolate a human MTase genomic
 CC clone (sequence given in 092813). Absence of MTase from a cell
 CC is indicative of malignancy.
 SQ Sequence 2763 BP; 778 A; 501 C; 602 G; 805 T;

Query Match 28.3%; Score 240; DB 15; Length 2763;
 Best Local Similarity 99.2%; Pred. No. 9,506-159;
 Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 aggtcttataagactgtctaaagactagactcggctgcctcctcaaggagcaatgg 1021

QY 570 AGGTCTTATAGAGACTGCTAAGAGCTGAGCTCGGTCCACTCAAGAGGACATGG 629
 |||||
 Db 1022 tcacaatcgaggagactcgttttaagctccggcgagaaagcttaagctcgcacctggg 1081
 |||||
 QY 630 TCACAAATCGAGGAGCTCTTTAGCTCCGGCGAAGAGCTTATGTTCCGACCTGGG 689
 |||||
 Db 1082 gggcgagatgtatcaacaatgacacagttccagagtggtcttctgtaaggagctggaa 1141
 |||||
 QY 690 GGGCGGATGTTATCAACATGACCACTCCAGAGTGTTCTTGTCTAAGAGGCTGGAA 749
 |||||
 Db 1142 ttgttacgcaagtatcgacatcgagcagagattatgactgtggaaggagcagaggaag 1201
 |||||
 QY 750 TTGTTCACGCAAGTATCGCATGCGGACAGATTATGACTGCTGGAAGAGACAGAGGAG 809
 |||||
 Db 1202 cagt 1205
 |||||
 QY 810 CAGT 813

RESULT 3
 ID 099202 standard; DNA; 2784 BP.

AC 099202;
 DT 07-MAR-1996 (first entry)
 DE Pseudomonas putida methylthioadenosine-phosphorylase DNA.
 KW Enzyme: MTase; L-Met-L-deamino-gamma-mercaptopethane-lyase;
 KW chemotherapy; cancer therapy; methionine starvation; ss.
 OS Pseudomonas putida.
 FH Key Location/Qualifiers
 FT CDS 1..2763
 FT /tag- a
 FN M09517908-A1.
 PD 06-JUL-1995.
 PR 22-DEC-1994; 014919.
 PR 29-DEC-1993; US-176413.
 PA (REGC) UNIV CALIFORNIA.
 PI Carson DA, Nobori T;
 WP1: 95-246192/32.
 PT Selective methionine starvation of methylthioadenosine
 PT phosphorylase negative tumour cells - used in chemotherapy of
 PT mammalian malignant cells.
 PS Disclosure: Page 27-28; 46pp; English.
 CC Antigenic peptides may be generated from this sequence encoding
 CC methyladenosine-phosphorylase, which are then used to generate
 CC antibodies specific for MTase. The produced antibodies may be
 CC used in an immunoassay for the detection of MTase.
 SQ Sequence 2784 BP; 781 A; 505 C; 598 G; 804 T;

Query Match 28.3%; Score 240; DB 16; Length 2784;
 Best Local Similarity 99.2%; Pred. No. 9,506-159;
 Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 aggtcttataagactgtctaaagactagactcggctgcctcctcaaggagcaatgg 1021
 |||||
 QY 570 AGGTCTTATAGAGACTGCTAAGAGCTGAGCTCGGTCCACTCAAGAGGACATGG 629
 |||||
 Db 1022 tcacaatcgaggagactcgttttaagctccggcgagaaagcttaagctcgcacctggg 1081
 |||||
 QY 630 TCACAAATCGAGGAGCTCTTTAGCTCCGGCGAAGAGCTTATGTTCCGACCTGGG 689
 |||||
 Db 1082 gggcgagatgtatcaacaatgacacagttccagagtggtcttctgtaaggagctggaa 1141
 |||||
 QY 690 GGGCGGATGTTATCAACATGACCACTCCAGAGTGTTCTTGTCTAAGAGGCTGGAA 749
 |||||
 Db 1142 ttgttacgcaagtatcgacatcgagcagagattatgactgtggaaggagcagaggaag 1201
 |||||
 QY 750 TTGTTCACGCAAGTATCGCATGCGGACAGATTATGACTGCTGGAAGAGACAGAGGAG 809
 |||||
 Db 1202 cagt 1205
 |||||
 QY 810 CAGT 813


```

RESULT 4
ID 010572 standard; DNA: 1047 BP.
AC 010572:
DE 09-APR-1991 (first entry)
DE Human Natriuretic Peptide Receptor B.
KM NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
OS hyperaldosteronism; glaucoma; guanylyl cyclase.
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label- signal sequence
FT Protein 12
FT /label- mature NPRB
FT Domain 23..455
FT /label- extracellular domain
FT /note- "binds natriuretic peptides A,B and C]"
FT Domain 456..456
FT /label- transmembrane domain
FT Domain 479..1047
FT /label- cytoplasmic domain
FT /note- "GC and protein kinase activity"
FT Modified-site 24..26
FT /label- N-glycos-site
FT Modified-site 35..37
FT /label- N-glycos-site
FT Modified-site 161..163
FT /label- N-glycos-site
FT Modified-site 195..197
FT /label- N-glycos-site
FT Modified-site 244..246
FT /label- N-glycos-site
FT Modified-site 277..279
FT /label- N-glycos-site
FT Modified-site 349..351
FT /label- N-glycos-site
FT Modified-site 600..602
FT /label- N-glycos-site
PN MO9100292-A.
PD 10-JAN-1991:
PD 22-JUN-1990: U03586.
PR 23-JUN-1989: US-370673.
PA (GETH ) GENENTECH INC.
PI Chang M, Goeddel D, Lowe D:
PI WPI: 91-036711/05.
DR N-PSDB: 010324.
DE Natriuretic protein receptor B - for diagnosis and treatment of
FT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS Claim 3; Fig 1: 49pp; English.
CC The sequence was derived from the DNA encoding natriuretic peptide
CC receptor B, NPRB, having guanylyl cyclase (GC) activity and protein
CC kinase activity. The DNA can be inserted into expression vectors
CC for the prodn. of the protein, opt. after being mutated to produce
CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr-
CC 114,952). The protein (or variants) can be used in treatment of
CC natriuretic peptide disorders, and also to isolate peptides using
CC affinity chromatography. Antibodies with affinity for NPRB can
CC also be prepd.
SQ Sequence 1047 BP: 87 A; 15 C; 83 G; 51 T;

Query Match 8.6%; Score 73; DB 2; Length 1047;
Best Local Similarity 8.3%; Pred. No. 2,34e-33;
Matches 66; Conservative 233; Mismatches 506; Indels 9; Gaps 9;
Db 221 grnyvncgnmmhnnnnnnnnnnngdyvnyndvngnsnragntztrvnmndnr 280
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 131 GGCACACACACACCGCGCTGAAGATGTGATGGAACAGCGCTGATGATCA 190
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 281 mnananranntvnyvtyrnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 340
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 191 GAAATTTTAAAGAGAACTGAAAAATATGTGATCTCATTTGGCAAGCATCTGAT 250
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 341 mnyvnnvnnnnnnngtrndgrnnvnmkmgrryhygtgnv-ymdkndntndvnmamd 399
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

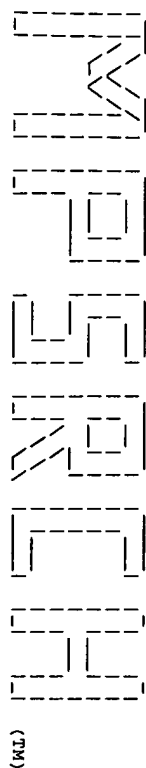
```

```

Oy 251 GCCTTAATTTGGGGAAGATAAAATGTGATTCGATCCCTCTCGAAGCATGGAAG 310
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 400 ndsgdnnaahys-ganknnwvtgrnnnwvxyannsdnnncandhdhscdktinstan 458
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 311 CAGCACACCATCATCTCTCAAGAGCTCAACACAGGCAACATCTGGCTTTGAAGAA 370
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 459 vanygnntnmngvsnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 518
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 371 GAGGCGCTGACACATGTCAATGACACACACAGCTTGTGCTTGTAGGAGAG-ATT-C 428
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 519 nrgssygsnmtahgkyunnantghknvankhvnkkrnnntrvnnnnkhrvnnnh 578
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 429 -AGCCCGG-CGATATGTGATATGATGATGATGATGATGATGATGATGATGAT 486
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 579 ntrnnnacdnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 638
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 487 GTCTTCTATGATGAGAGCATCTTGTGCCAGAGAGAGTGCATATTCGAATGCTGA 546
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 639 nshgnsksnncvsvsnvsnvsnvsnvsnvsnvsnvsnvsnvsnvsnvsnvsnv 698
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 547 GCCGTTTGGCCCCAAAGAGAGAGAGGTTCTTATAGAGAGCTGCTAGAGAGAGAG 606
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 699 tlymaadvysgnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 758
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 607 GTGCCACTCAAGAGGAGACATGTCACATGACGAGGAGCTGTTAGCTCCGGGACGA 666
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 759 nnnvnnnn-rcwandnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 817
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 667 AAGCTTCATGTTCCGACCTGGGGGGGAGATGTATCAACATGACACAGTCCAGAGGT 726
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 818 nrtnnnnnnkrrannnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 876
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 727 GGTCTTGTGAGAGAGGCTGGAATTTGTACGACAGTATCCGATGAGGACGATTATGA 786
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 877 stnnnnvntnnndnycndannndndndvkv-vntngdaymvsngngnngnrrhnnann 935
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 787 CTGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 936 andavssnnrnrhndnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 995
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 847 AAGCTTAATAAGCAAAAGCTTACTGCTACTGCTACTGCTACTGCTACTGCTACTG 906
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 996 ankhnvsettkdaandngcnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 1031
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 907 ATGTCAGAAACCTTCATTAACCTGAGAAATATGTC 942
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 5
ID 010572 standard; DNA: 1047 BP.
AC 010572:
DE 09-APR-1991 (first entry)
DE Human Natriuretic Peptide Receptor B.
KM NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
OS hyperaldosteronism; glaucoma; guanylyl cyclase.
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label- signal sequence
FT Protein 12
FT /label- mature NPRB
FT Domain 23..455
FT /label- extracellular domain
FT /note- "binds natriuretic peptides A,B and C]"
FT Domain 456..456
FT /label- transmembrane domain
FT Domain 479..1047
FT /label- cytoplasmic domain
FT /note- "GC and protein kinase activity"
FT Modified-site 24..26
FT /label- N-glycos-site
FT Modified-site 35..37
FT /label- N-glycos-site
FT Modified-site 161..163

```

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelligentGenetics, Inc.

Msearch_n n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Feb 10 23:56:55 1997; MasPar time 440.34 Seconds
850.327 Million cell updates/sec
Tabular output not generated.

Title: >US-08-674-311-1
Description: (122-970) from US08674311.seq (2 of 2)
Perfect Score: 849
N.A. Sequence: 122 ATGGCCCTCTGGCACCACAC.....CTGTTTATTACAGACAT 970
Comp: TACCGGAGACCGTGTGTGTG.....GACAAATAATAGTTCTGTA

Scoring table: TABLE default

Nmatch STD : Dbase 0; Query 0

Searched: 630489 seqs, 220513910 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99

EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174
175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
195:EST195 196:EST196 197:EST197 198:EST198 199:EST199
200:EST200

173:enSTS1 174:enSTS2

Statistics: Mean 10.783; Variance 1.823; scale 5.916
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	67	7.9	168	50	HUMNK726	Human keratinocyte CD	1.79e-90
2	49	5.8	400	134	G17388	human STR SHGC-14428	9.95e-53
3	49	5.8	400	173	HS388293	human STR SHGC-14428	9.95e-53
4	24	2.8	447	2	CELK008E4F	C.elegans CDNA clone	8.11e-08
5	23	2.7	321	1	ARTS2467	A. thaliana transcrib	2.15e-06
6	23	2.7	341	2	CELK002F5F	C.elegans CDNA clone	2.15e-06
7	23	2.7	435	107	T22905	4913 Arabidopsis thal	2.15e-06
8	22	2.6	332	4	CELK062D9R	C.elegans CDNA clone	5.08e-05
9	22	2.6	981	150	W13162	me74d12.r1 Soares mou	5.08e-05
10	22	2.6	981	159	MM1627	me74d12.r1 Soares mou	5.08e-05
11	21	2.5	202	168	HSW04230	z433e02.r1 Soares fet	1.06e-03
12	21	2.5	202	147	W04230	z433e02.r1 Soares fet	1.06e-03
13	21	2.5	301	114	T47192	yf53e11.r1 Homo sapie	1.06e-03
14	21	2.5	337	83	R41555	yf88c05.r1 Homo sapie	1.06e-03
15	21	2.5	336	44	HSC3AD011	H. sapiens partial CD	1.06e-03
16	21	2.5	342	151	W17604	mb74h08.r1 Soares mou	1.06e-03
17	21	2.5	342	151	MM6049	mb74h08.r1 Soares mou	1.06e-03
18	21	2.5	349	135	HS273ZC9	H. sapiens (D125336)	1.06e-03
19	21	2.5	357	98	R30606	16961 Arabidopsis tha	1.06e-03
20	21	2.5	380	22	H52428	yq78c10.s1 Homo sapie	1.06e-03
21	21	2.5	394	72	R02943	PK02f06.r2 Caenorhabd	1.06e-03
22	21	2.5	429	74	R08098	yf17902.r1 Homo sapie	1.06e-03
23	21	2.5	446	117	T58479	yf61d12.r1 Homo sapie	1.06e-03
24	21	2.5	454	54	N25437	yx37a12.r1 Homo sapie	1.06e-03
25	21	2.5	465	72	R03064	PK3e07.s2 Caenorhabd	1.06e-03
26	21	2.5	471	75	R13815	yf1f08.r1 Homo sapie	1.06e-03
27	21	2.5	586	54	N25469	yx38b11.r1 Homo sapie	1.06e-03
28	21	2.5	631	109	T30602	EST109351 Homo sapiens	1.06e-03
29	21	2.5	958	148	W08053	mb38f09.r1 Soares mou	1.06e-03
30	21	2.5	958	165	MM0531	mb38f09.r1 Soares mou	1.06e-03
31	21	2.5	970	169	MM2227	me83d08.r1 Soares mou	1.06e-03
32	21	2.5	970	169	W13222	me83d08.r1 Soares mou	1.06e-03
33	20	2.4	205	45	HSCZD092	H. sapiens partial CD	1.93e-02
34	20	2.4	332	104	T07338	EST05628 Homo sapiens	1.93e-02
35	20	2.4	349	115	T53348	yf88c10.r1 Homo sapie	1.93e-02
36	20	2.4	365	81	R34431	yf85f08.r1 Homo sapie	1.93e-02
37	20	2.4	381	102	R1C52672A	Rice CDNA, partial se	1.93e-02
38	20	2.4	405	17	H51172	EST107775 Rattus sp.	1.93e-02
39	20	2.4	420	167	HS990322	yf71e05.r1 Homo sapie	1.93e-02
40	20	2.4	434	14	H27151	yf19e06.r1 Homo sapie	1.93e-02
41	20	2.4	435	9	H09657	yf19e06.r1 Homo sapie	1.93e-02
42	20	2.4	442	28	H71681	ys01c07.r1 Homo sapie	1.93e-02
43	20	2.4	450	159	HS220331	z213c08.r1 Soares fet	1.93e-02
44	20	2.4	489	90	R64583	yf17d01.s1 Homo sapie	1.93e-02
45	20	2.4	548	146	W01585	z280f09.r1 Soares fet	1.93e-02

ALIGNMENTS

RESULT 1
LOCUS HUMNK726 168 bp mRNA
DEFINITION Human keratinocyte CDNA, clone 726.
ACCESSION D29573
NID 959506
KEYWORDS EST(expressed sequence tag).
SOURCE Homo sapiens Epidermis
ORGANISM Homo sapiens
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
REFERENCE 1 (bases 1 to 168)

AUTHORS Yamashiki, K., Morishima, Y., Ueda, E., Nonomura, K., Kibe, S.,
 Yamashiki, K. and Yasuno, H.
JOURNAL Unpublished (1994)
REFERENCE 2 (sites)
AUTHORS Konishi, K., Morishima, Y., Ueda, E., Kibe, Y., Nonomura, K.,
 Yamashiki, K. and Yasuno, H.
TITLE Characterization of the genes expressed in human keratinocytes: analysis
 of 607 randomly isolated cDNA sequences
JOURNAL Biochem. Biophys. Res. Commun. 202 (2), 976-983 (1994)
MEDLINE 94324994
COMMENT Submitted (11-Mar-1994) to DDBJ by:
 Kiyofumi Yamashiki
 Department of Dermatology
 Kyoto Prefectural University of Medicine
 Kamigyo-ku
 Kyoto, Kyoto, 602
 Japan
 Phone: 75-251-5587
 Fax : 75-251-5587
FEATURES PROJECT="human epidermal keratinocyte_ESTS",
 location/Qualifiers
 1..168
 /organism="Homo sapiens"
 /cell_line="Primary culture"
 /cell_type="Keratinocyte"
 /sequenced_mol="cDNA to mRNA"
 /tissue_type="Epidermis" 62 t
BASE COUNT 43 a 26 c 37 g
ORIGIN
 Query Match 7.9%; Score 67; DB 50; Length 168;
 Best Local Similarity 93.3%; Pred. No. 1,79e-90;
 Matches 83; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
 Db 19 agcttggaataat-gattggaacagccttggaatccagaaatagtagaagaagactt 77
 |||||
 Oy 153 AGATTGGAATTAATGTGTGGAAAGCGCTGATGATTCAGAAATTTAGAGGAAGAACTG 212
 |||||
 Db 78 ttaaatatgtgatcac-ccattgtggaag 105
 |||||
 Oy 213 AAAAATATGTGATCTCATTGGCAG 241
 |||||
RESULT 2
LOCUS G17388 400 bp DNA STS 05-MAR-1996
DEFINITION human STS SHGC-14428 clone pg-7320.
ACCESSION G17388
NTD g1214814
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human Plasmid clones, generated from a lymphoblastoid cell line
 from a human male. Localized to human chromosome 9 by analysis on
 the NIGMS Human/Rodent Somatic Cell Hybrid Panel #1. Coriell
 Institute for Medical Research, Camden, NJ 08103.
ORGANISM Homo sapiens
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrates; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
 Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates
 Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 400)
AUTHORS Myers, R.M.
JOURNAL Unpublished (1996)
COMMENT
 Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myerseshgc.stanford.edu
PRIMER A: TGTGGATCTCCATTGGCA
 Primer B: TCACCAATCAGCGCAGAG
STS size: 90

```

PCR Profile:
Initial Incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

FEATURES
source Chromosome 9.
Location/Qualifiers
1..400
/organism="Homo sapiens"
/note="human"
61..150
/map="g"
61..80
/map="g"
primer_bind /map="g"
primer_bind complement(131..150)
BASE COUNT 95 a 74 c 83 g 101 t 47 others
ORIGIN

Query Match 5.8%; Score 49; DB 134; Length 400;
Best Local Similarity 98.2%; Pred. No. 9,95e-53;
Matches 55; Conservative 0; Mismatches 0; Indels 1; Gaps 1

Db
28 atccagaattttagaagaaga-a-citgaataatgtgtgatactccattgccaag 82
|||||
186 ATCCAGAAATTTAGAGAGAGAACTGAAAATATGTGTGATCTCCATTGGCAG 241

RESULT 3 standard; DNA; STS: 400 BP.
ID HS388293 standard; DNA; STS: 400 BP.
AC G17386;
DT 07-MAR-1996 (Rel. 4.7, Created)
DT 07-MAR-1996 (Rel. 4.7, Last updated, Version 1)
DE human STS SHGC-14428 clone pg-7320.
KW primer: sequence tagged site; STS sequence.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
OC [1]
RN RP 1-400
RA Myers R.M.;
RT ;
RU Unpublished.
CC Contact: Richard M. Myers Stanford Human Genome Center (SHGC)
CC Stanford University School of Medicine Department of Genetics,
CC M-344, Stanford, CA 94305, USA Tel: 4157259687 Fax: 4157259689
CC Email: myers@shgc.stanford.edu Primer A: TGTGATATCTCCATTGGCA
CC Primer B: TACCAATACAGCCAGCAG STS size: 90 PCR profile: Initial
CC incubation: 94 degrees C for 90 seconds Denaturation: 94 degrees C
CC for 15 seconds Annealing: 62 degrees C for 23 seconds
CC Polymerization: 72 degrees C for 30 seconds PCR Cycles: 30 Thermal
CC Cycler: Perkin Elmer 9600 Protocol: Template: 25 ng Primer: each 1
CC uM dNTPs: each 200 uM Tag Polymerase: 0.05 units/ul Total Vol: 10
CC ul Buffer: MgCl2: 2.5 mM KCl: 50 mM Tris-HCl: 20 mM pH: 8.3
CC Chromosome 9, NCBI g1.121414
CC Location/Qualifiers
FH

```



```

FT source 1..400
FT /organism="Homo sapiens"
FT /note="human"
FT STS 61..150
FT /map="9"
FT primer_bind 61..80
FT /map="9"
FT primer_bind complement(131..150)
FT /map="9"
SQ Sequence 400 BP; 95 A; 74 C; 83 G; 101 T; 47 other;

Query Match 5.8%; Score 49; DB 173; Length 400;
Best Local Similarity 98.2%; Pred. No. 9,95e-53;
Matches 55; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 28 atccgaatttagaagaaga-ctgaaatattgatactcattggcaag 82
|||||
0y 186 ATCCGAATTTAGAGAGAGACTGAAATATGTGATCTCATTTGGCAAG 241

RESULT 4 CELK008E4F 447 bp mRNA EST 07-AUG-1994
LOCUS DEFINITION C.elegans cDNA clone yK8e4 : 5' end, single read.
ACCESSION D37705
NID 9526087
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain CB1489 him-8(e1489)) (library: Yuji Kohara unpublished cDNA) Hermaphrodite, male varied whole animal cDNA to mRNA.
ORGANISM Caenorhabditis elegans
Eukaryota; Animalia; Eumetazoa; Nematoda; Secernentea; Rhabditida;
Rhabditina; Rhabditidae; Rhabditidae.
REFERENCE 1 (bases 1 to 447)
AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and Tabara,H.
TITLE Toward an expression map of the C.elegans genome(5'end)
JOURNAL Unpublished (1994)
COMMENT Submitted (1-Jul-1994) to DDBJ by: Yuji Kohara
National Institute of Genetics
Gene Library Lab
Yata 1111, Mishima
Shizuoka 411
Phone: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.jp.
Location/Qualifiers

FEATURES
source 1..447
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/dev_stage="varied"
/sequenced_mol="cDNA to mRNA"
/sex="Hermaphrodite, male"
/tissue_type="whole animal"
/clone_lib="Yuji Kohara unpublished cDNA"

BASE COUNT 153 a 114 c 77 g 103 t
ORIGIN

Query Match 2.8%; Score 24; DB 2; Length 447;
Best Local Similarity 80.0%; Pred. No. 8.11e-08;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 113 atgtgataatccattggcaatccatccacgacatgat 152
|||||
0y 219 ATGTGATACTCCATTGGCAAGCCATCTGATGCCTTAAT 258

RESULT 5 ATTS2467 321 bp RNA EST 14-MAR-1994
LOCUS DEFINITION A. thaliana transcribed sequence; clone TA1332; 3' end; Similar to
ACCESSION ribosomal protein L7a; Oryza sativa.
NID 230721
KEYWORDS expressed sequence tag; partial cDNA sequence.

```

```

SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Eukaryota; Eukaryota; Chlorophyta/Embryophyta
group; Charophyta/Embryophyta group; Embryophyta; Magnoliopsida;
Magnoliopsida; Dilleniidae; Caprariales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 321)
AUTHORS Phillips,G. and Gigot,C.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1994) to the EMBL/GenBank/DBJ databases. CNRS,
GDR-1003 ACS, INRA, Laboratoire de Biologie Moléculaire, BP 27,
31326 Castanet-Tolosan cedex, France.
E-mail: gdt-sv@toulouse.inra.fr. On behalf of: Laboratoire de
Biologie Moléculaire des Plantes - CNRS, Gigot Claude / L512, 12
Rue du Général Zimmer, 67084 Strasbourg Cedex, France.
E-mail: philip@scilla.u-strasbg.fr
2 (bases 1 to 321)
CNRS.

REFERENCE The Arabidopsis thaliana transcribed genome: the GDR cDNA program
AUTHORS Unpublished
TITLE Cloning vector: Lambda ZAP11;
JOURNAL Physiological condition: cycling cells.
COMMENT full automatic.
Similarity detected by blastx against GB entry D12631.

FEATURES
source 1..321
/organism="Arabidopsis thaliana"
/clone="TA1332"
/tissue_type="cell suspension culture of ecotype columbia"
/clone_lib="AC16H"
complement(188..>321)
/codon_start=3
/product="ribosomal protein"
/db_xref="PID:9460355"
/translaction="LTSYRVEDRRKKGSGIMSKSQAKTKAKERYIAKQBMN"

BASE COUNT 99 a 68 c 51 g 103 t
ORIGIN

Query Match 2.7%; Score 23; DB 1; Length 321;
Best Local Similarity 78.0%; Pred. No. 2.15e-06;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 205 cagctccttagcataaccctcctccttagctgtctg 245
|||||
Cp 746 CAGCCTCTTAGCAAGACCACTCTGAGACGTGTCATG 706

RESULT 6 CELK002F5F 341 bp mRNA EST 06-AUG-1994
LOCUS DEFINITION C.elegans cDNA clone yK2f5 : 5' end, single read.
ACCESSION D36187
NID 9525402
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain CB1489 him-8(e1489)) (library: Yuji Kohara unpublished cDNA) Hermaphrodite, male varied whole animal cDNA to mRNA.
ORGANISM Caenorhabditis elegans
Eukaryota; Animalia; Eumetazoa; Nematoda; Secernentea; Rhabditida;
Rhabditina; Rhabditidae; Rhabditidae.
REFERENCE 1 (bases 1 to 341)
AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and Tabara,H.
TITLE Toward an expression map of the C.elegans genome(5'end)
JOURNAL Unpublished (1994)
COMMENT Submitted (1-Jul-1994) to DDBJ by: Yuji Kohara
National Institute of Genetics
Gene Library Lab
Yata 1111, Mishima
Shizuoka 411
Phone: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.jp.
Location/Qualifiers

FEATURES
source 1..341

```

[illegible]

	7					
RESULT						
LOCUS	T22905	435 bp	mRNA	EST		28-MAR-1995
DEFINITION	4913 Arabidopsis thaliana cDNA clone 108B10T7.					
ACCESSION	T22905					
NID	9504346					
KEYWORDS	EST					
SOURCE	thale cress clone=108B10T7 library=Lambda-PRL2 strain-var col accession=U01843 protein=beta1 est1=beta2-not Lambda PRL2 is a					

ES3 creess clone-10BB1071 library-lambda-PR2 strain-var columbia thale cress clone-10BB1071 library-lambda-PR2 strain-var columbia vector-lambda zip-Lox Raitel-Sal Raitel-Not lambda PR2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-roseettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BR1's lambda zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA.

ORGANISM	REFERENCE	TITLE	JOURNAL
Arabidopsis thaliana	Neuman, T., de Bruijn, F. J., Green, P., Keegstra, K., Kende, H., McInosh, L., Ohlroge, J., Ralthe, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.	Large-scale partial sequencing of anonymous Arabidopsis cDNA clones	Plant Physiol. 106, 1241-1255 (1994)
Eucaryotae; Embryophyta; Magnoliophyta; Magnoliopsida; Caprales; Brassicaceae; Arabidopsys.	1 (bases 1 to 435)		

CONTACT: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313c@doe.cm.msu.edu.
Location/Qualifiers

FEATURES	source	Location/Qualifiers
1.435		/orgnslsm="Arabidopsis thaliana"
		/clone="108B107"
		/strain="var Columbia"
		/note="thale cress"
114	a	86 c 112 g 115 t
BASE COUNT		8 others
ORIGIN		

Query Match	2.76;	Score 23;	DB 107;	Length 435;
Best Local Similarity	78.08;	Pred. No. 2.15e-06;		
Matches	32;	Conservative	0;	Mismatches 9;
			Indels	0;
			Gaps	0;

Db 294 caagaccaaagctaaagagagaggtattctgaagagctg 334
 || ||||| || ||||| ||||| ||||| ||||| |||||
 QY 706 CATGACCACTTCAGAGGTGGTCTTGCTAAGAGGCTG 746

RESULT	8
LOCUS	CELK062D9R 332 bp mRNA EST 09-DEC-1995

DEFINITION	C. elegans CDNA clone yK62d3 : 3' end, single read.
ACCESSION	D71125
NID	9110834
KEYWORDS	EST/repressed sequence tag.
SOURCE	Uncollected C. elegans (strain N2), (library: yuji Kohara uncollected cDNA) Hemaphrophidite embryo CDNA to mRNA.

REFERENCE	TITLE	JOURNAL	COMMENT
1 (bases 1 to 333)			
Kohara, Y., Mochizuki, T., Tabara, H., Sugimoto, A., Watanabe, H. and Nishigaki, A.	Toward an expression map of the <i>C. elegans</i> genome	Unpublished (1995)	Submitted (23-Aug-1995) to DBJ by: Yutai Kohara

Yoshitaka Kohara
Genetic Laboratory
National Institute of Genetic
Yata 1-11, Mishima Shizuoka
411 Japan
Phone: 0559-81-6854
Fax : 0559-81-6855
Email:ykohara@dbj.nig.ac.jp

```

FEATURES
source
Location/Qualifiers
1..332
/corantum="Caenorhabditis elegans"

```

```

/c1one_lib="yuj1 Kohara unpublished cDNA"
BASE COUNT      130 a      46 c      75 g      81 t
ORIGIN

```

Query Match	2.68;	Score 22;	DB 4;	Length 332;
Best Local Similarity	92.30;	Pred. No. 5.08e-05;		
Matches	24;	Conservative	0;	Mismatches 2;
			Indels 0;	Gaps 0

Db	193	ataaaacagaaaaatggcgatat	218
Cp	960	ATAAAACAGAAAAC	935
		CTGGCCATATT	

	9	RESULT						
LOCUS	wj3162	981 bp	mRNA	EST	26-APR-1996			
DEFINITION	ma74d12.r1 Soares mouse pNMf19.5 Mus musculus cDNA 5' similar to SW:CD4_MOUSE P30285 CELL DIVISION PROTEIN KINASE 4 ; .							
ACCESSION	wj3162							
NID	g1287199							

ORGANISM Mus musculus
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
1 (bases 1 to 981)
Matta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Hawthorne, J., Jones, A., Kohn, M., Lander, E., Linton, J., Marra, M.,
McEwen, R., Miller, C., Morris, M., Nadeau, J., Naylor, J., Nizenz, L.,

TITLE	JOURNAL	COMMENT
The WashU-HHMI Mouse EST Project	Unpublished (1996)	

Contact: Tarrja M/Mouse EST Project
WashU-PHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1600
Tel: 314 286 1810
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: ERPrimer
High quality sequence stop: 258.

NCBI g1: 1287199

FEATURES
source
Location/Qualifiers

1..981

/organism="Mus musculus"
/note="Vector: p773D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCACTGAGTGGAGCGCGCCATTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p773 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo."
/clone_lib="Soares mouse p33MF19.5"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>981

BASE COUNT 203 a 227 c 304 g 236 t 11 others

ORIGIN

Query Match 2.6%; Score 22; DB 150; Length 981;
Best Local Similarity 86.5%; Pred. No. 5.08e-05;

Matches 32; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Db 328 cacagtcgtgagtgagc-ctgttaagaagagctga 363

Oy 712 CACAGTTCGAGGAGTGTTCTGTACGAGGCTGGA 748

RESULT 10
ID MM1627 standard; RNA; EST; 981 BP.

AC W13162;

DT 30-APR-1996 (Rel. 47, Created)

DE ma74dl.r1 Soares mouse p33MF19.5 Mus musculus cDNA 5' similar to

DE SW:CDK4_MOUSE P30285 CELL DIVISION PROTEIN KINASE 4 ;

KW EST.

OS Mus musculus (mouse)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.

RN [1]

RP 1-981

RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,

RA Dubuque T., Geisels S., Kucaba T., Lacy M., Le M., Martin J.,

RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,

RA Moore B., Theisling B., Wylie T., Lennon G., Soares B., Wilson R.,

RA Waterston R.;

RA "The Washu-HHMI Mouse EST Project";

RA Unpublished.

CC Contract: Marra M/Mouse EST Project Washu-HHMI Mouse EST Project

CC Washington University School of Medicine 4444 Forest Park Parkway,

CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810

CC Email: mouseest@wustl.edu This clone is available

CC (info@image.lnl.gov) for further information. Seq primer: ERPrimer

CC High quality sequence stop: 258. NCBI g1: 1287199

FH Key Location/Qualifiers

FH 1..981

FT /organism="Mus musculus"

FT /note="Vector: p773D (Pharmacia) with a modified

FT polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

FT was primed with a Not I - oligo(dT) primer [5'

FT TGTTACCACTGAGTGGAGCGCGCCATTTTCTTTTCTTTT 3'],

FT double-stranded cDNA was size selected, ligated to Eco RI

FT adapters (Pharmacia), digested with Not I and cloned into

FT the Not I and Eco RI sites of a modified p773 vector

(Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo."
/clone_lib="Soares mouse p33MF19.5"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>981

SQ Sequence 981 BP; 203 A; 227 C; 304 G; 236 T; 11 other:

Query Match 2.6%; Score 22; DB 169; Length 981;

Best Local Similarity 86.5%; Pred. No. 5.08e-05;

Matches 32; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Db 328 cacagtcgtgagtgagc-ctgttaagaagagctga 363

Oy 712 CACAGTTCGAGGAGTGTTCTGTACGAGGCTGGA 748

RESULT 11
ID HSW04230 standard; RNA; EST; 202 BP.

AC W04230;

DT 30-APR-1996 (Rel. 47, Created)

DE za43e02.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone

DE 295322 5'.

KW EST.

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homidae.

RN [1]

RP 1-202

RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,

RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,

RA Trevasakis E., Waterston R., Williamson A., Wohldmann P., Wilson R.,

RA "The Washu-Merck EST Project";

RA Unpublished.

CC Contract: Wilson RK Washu-Merck EST Project Washington University

CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,

CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

CC est@wustl.edu This clone is available royalty-free through

CC LNL; contact the IMAGE Consortium (info@image.lnl.gov) for

CC further information. Seq primer: mob.REGA-ET High quality sequence

CC stop: 176. NCBI g1: 1276338

FH Key Location/Qualifiers

FH 1..202

FT /organism="Homo sapiens"

FT /note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)

FT with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;

FT 1st strand cDNA was primed with a Pac I - oligo(dT) primer

FT [5' AACGCGAAGATTAATTAAGACATCTTTTCTTTTCTTTT 3'],

FT double-stranded cDNA was ligated to Eco RI adapters

FT (Pharmacia), digested with Pac I and cloned into the Pac I

FT and Eco RI sites of the modified p773 vector. Library was

FT through one round of normalization. Library constructed by

FT Bento Soares and M.Fatima Bonaldo."

FT /clone_lib="295322"

FT /clone_lib="Soares fetal liver spleen INFLS"

FT /sex="male"

FT /dev_stage="20 week-post conception fetus"

FT /lab_host="DH10B (ampicillin resistant)"

FT <1..>202

FH mRNA

SQ Sequence 202 BP; 80 A; 33 C; 43 G; 44 T; 2 other:

Query Match 2.5%; Score 21; DB 168; Length 202;

Best Local Similarity 88.9%; Pred. No. 1.06e-03;

Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 19 gacatacctgagtgagctagtagagc 45

Oy 900 GACCTATCTGAGGTAGTGTAGTACGAC 874

Tue Feb 11 11:27:39 1997

US-08-674-311-1-02.rst

Page 6

RESULT	12
LOCUS	M04230
DEFINITION	za496c2.r1 Soares fetal liver spleen lncfIs Homo sapiens cDNA clone 295322 5'.
ACCESSION	M04230
NID	g1276338
KEYWORDS	EST.
SOURCE	human. human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryotae: mitochondrial eukaryotes; Metazoa: Chordata: Vertebrata: Eutheria; Primates: Catarrhini: Homnidae: Homo. 1 (bases 1 to 202)
AUTHORS	Hillier,L., Clark,M., Dubugue,T., Ellison,K., Hawkins,M., Holman,M., Hultman,M., Kiedrzyk,T., Le,M., Lennon,G., Marr,M., Parsons,J., Nikin,L., Kniffling,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE	The WashU-Merck EST Project
JOURNAL	
COMMENT	Unpublished (1995)
Contact:	wilson@x

Contact: Wilson RK
Washu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mob.RBG4-51
High quality sequence stop: 176.

NCBI g1: 1276338

FEATURES
SOURCE

```

/organism="Homo sapiens"
/organ="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polynuker; Site.1: Pac I; Site.2: Eco RI;
1st strand cDNA was primed with a Pac I - 01190(dT) primer
15' AACGCGAGCAATTAATTAAGACTCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Felma Bonaldo."
/clone="293122"
/clone_11b="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
<1..>202

```

Query Match	2.58;	Score 21;	DB 147;	Length 202;
Best Local	Similarity 88.9%;	Pred. No. 1.06e-03;		
Matches	24;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;

```

Db      19  gaccatacctgaagtgatgtgagc 45
        |||||
Cp      900 GACCCTATCTGAGGTATGGTAGTGAGC 87

```

RESULT	LOCUS	DEFINITION	EST	08-FEB-1995
13	T47192	301 bp	MRNA	
	y53ell.r1	Homo sapiens CDNA clone 74924 5'	similar to gb:U11566	
		60S RIBOSOMAL PROTEIN L18 (HUMAN).		

ACCESSION	T4/192
NID	9649174
KEYWORDS	EST.
SOURCE	human clone=74924 library=Stratagene ovary (#937217)

vector-bluescriptript SK host-SOLR cells (kanamycin resistant).
primer-M3R1 Rstlet-EORI Rstlet-XhoI Cloned unidirectionally.
primer: Oligo dt. Total ovary tissue, normal 49 year old caucasian
female. Average insert size: 0.8 kb; Uni-ZAP XR Vector; 5' adaptor
sequence: 5'-GAAATTCGCGACAG-3'; 3' adaptor sequence:
5'-CTCGAGTTTTTTTTTTTTTT-3'.
ORGANISM
Homo sapiens
Eucaryote; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;

ORGANISM	Homo sapiens; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eucaryotes
REFERENCE	1 (bases 1 to 301)
AUTHORS	Hallier, B., Clark, N., Dubouche, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kuwabara, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, B., Rohlfing, T., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.
TITLE	Mashu-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	Other ESTs: yb3cell.s1
	Contact: Wilson RK

WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 1294
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@lml.lln.gov) for further information.

FEATURES	Location/Qualifiers
SOURCE	1..301
	/orcanitem="HOMO sapiens"

BASE COUNT	64 a	71 c	81 g	72 t	13 others
ORIGIN					

Query Match	2.5%	Score 21	DB 114	Length 301
Best Local Similarity	68.6%	Pred. No. 1.00e-03		
Matches	35	Conservative	0	Mismatches 16
			Indels	0
			Gaps	0

```
Dβ      155   cggnagatgaaggcttcctcgtgycgcggtgaacaacagangyccgltgttgyt    205  
          ||| | | | | | | | | | | | | | | | | | | | |  
CP      681   CGGACATGAAGCTTTCGCCCGGAGCTAAACGAGGTCCCTCGATTGTG    731
```

RESULT	14				
LOCUS	R41555	327 bp	mRNA	EST	22-MAY-1995
DEFINITION	yf88c05.s1 Homo sapiens CDNA clone 29736 3'.				
ACCESSION	R41555				
NID	g816856				
KEYWORDS	EST.				
SOURCE	human clone=29736 library=soares infant brain NIB vector=Lafmid BA				
	g816856 strain=sestret primer=promega -2lm3j Rsfle1-not				

Hind III-phosphorylated Not I whole brain from a 73 days postnatal female, 1st round cDNA was ligated into the Hind III-digested pTZ19 vector. The second round cDNA was primed with a Not I - Oligo(dT) primer [5'-ACCTGAGACATCCCGCCGCCAGCAATTATTTTATTTTTTTT-3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the lacZmid BA vector. Library went through one round of cDNA normalization. Library constructed by Bento Soares and M.Fátima Bonaldo.

ORGANISM

ORGANISM
Homo sapiens
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata.
Neurostomata; Chordata; Vertebrata; Gnathostomata; Osteichthyes
Sarcopterygii; Chonmata; Tetrapoda; Amniota; Mammalia; Theria.
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 327)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Hawthorn, J., Jones, R. W., Johnson, G., Marra, M.,
Mason, D., Miller, R., Mitchell, J., Moseley, J.,
Muller, R., Naylor, J., O'Brien, A., O'Brien, J.,
O'Brien, K., O'Brien, L., O'Brien, M., O'Brien, N.,
O'Brien, P., O'Brien, R., O'Brien, S., O'Brien, T.,
O'Brien, U., O'Brien, V., O'Brien, W., O'Brien, X,
O'Brien, Y., O'Brien, Z.

Holman, M., Hultman, M., Kucaba, T., Le, M., Lemmon, C., Miller, J., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Williamson, A.

TITLE The Washu-Merck EST Project

JOURNAL Unpublished (1995)
COMMENT

GDB: G00-402-083
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: estevatson.wustl.edu
High quality sequence stops: 127
Source: IMAGE Consortium, LNLN
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Location/Qualifiers

FEATURES
Source 1..327
/organism="Homo sapiens"
/clone="29736"
/note="human"

BASE COUNT 73 a 69 c 80 g 100 t 5 others
ORIGIN

Query Match 2.5%; Score 21; DB 83; Length 327;
Best Local Similarity 80.0%; Pred. No. 1.06e-03;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 51 aaattactaagaagacctgagcaaatgtgat 85
||||| ||||||| ||| |||||||
Gy 192 AAATTTAGAAGAGAGACTGMAAATATGTGAT 226

RESULT 15
LOCUS HSC3AD011 336 bp RNA EST 21-SEP-1995
DEFINITION H. sapiens partial cDNA sequence; clone c-3ad01.
ACCESSION F12463
F12463
NID 6708455
KEYWORDS partial cDNA sequence; transcribed sequence fragment.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homindae; Homo.

REFERENCE
TITLE Direct Submission
AUTHORS Submitted (19-JAN-1995) Genethon, B.P. 60, 91002 Evry Cedex France
JOURNAL B.P. 8, 94801 Villejuif Cedex France. E-mail: genexpress@genethon.fr
1 (bases 1 to 336)

REFERENCE
TITLE 2 (bases 1 to 336)
AUTHORS Genexpress.
JOURNAL The Genexpress cDNA program
COMMENT Unpublished

REFERENCE
TITLE 3 (bases 1 to 336)
AUTHORS Auffray, C., Behar, G., Bols, F., Bouchier, C., da Silva, C.,
Devignes, M.D., Duprat, S., Hougatte, R., Jumeau, M.N., Lamy, B.,
Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pletu, G., Pouliot, Y.,
Sebastiani-Kabakchis, C. and Tessier, A.
JOURNAL IMAGE: molecular integration of the analysis of the human genome
COMMENT and its expression

TITLE
JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
MEDLINE 95277534
COMMENT Cloning method: total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the lacmid BA
vector;

Sequencing method: single read, full automatic;

Primer: M13 reverse

cDNA sequence colinear to mRNA

Stretch removed: nothing

Normalization method: Bento Soares, P.N.A.S. 91:9228-9232(1994);

Genexpress library_id: C;

Genexpress sequence_id: Y1c-3ad01.

FEATURES
Source 1..336
Location/Qualifiers

/organism="Homo sapiens"
/isolate="muscular atrophy patient"
/clone_lib="normalized infant brain cDNA from B.Souares,
Psychiatry Dept. Columbia University USA"
/sex="female"

/tissue-type="total brain"
/dev-stage="3 months old"

BASE COUNT 82 a 78 c 103 g 68 t 5 others
ORIGIN

Query Match 2.5%; Score 21; DB 44; Length 336;
Best Local Similarity 85.7%; Pred. No. 1.06e-03;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 212 tucggaagatgaagcttcctgcccggga 239
||||| ||||||| ||| |||||||
Cp 683 TGCGACATGACGCTTCTCCCGGA 656

Search completed: Tue Feb 11 00:04:32 1997
Job time : 457 secs.

BEST AVAILABLE COPY

(TM)

Distribution rights by Intellegenetics, Inc

TACCGAGACCGTGGTGTG.....GACAAATAATGGTCTGTA

Gap 6

Dbase 0; Query 0

50278 seqs, 13829960 bases x 2

Listing first 45 summaries

n-issued

10:PCT94 11:PCT95 12:PCT96

Mean 8.502; Variance 4.199; scale 2.025

ived by analysis of the total score distribution.

SUMMARIES

2.4	3819	4	US-08-002-	Sequence 3, Applcattio	1.28e+00
-----	------	---	------------	------------------------	----------

	Sequence	/	Application
43	2.1	14311	1.43e+01
16	2.1	14311	1.43e+01

ALIGNMENTS

IMMEDIATE SOURCE:

Tue Feb 11 11:27:33 1997

US-08-674-311-1-02.rni

Page 2

```

CC CLONE: methyladenosine phosphatase
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 bp: 780 A; 505 C; 598 G; 803 T; 77 other;
SQ Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match 28.3%; Score 240; DB 10; Length 2763;
Best Local Similarity 99.28; Pctd. No. 9,33e-177;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 AGGTTCTTAGAGACTGCTAAGAGTACAGACTCCGGTCCACTCAAGGGGACAAATGG 1021
|||
OY 570 AGGTTCTTAGAGACTGCTAAGAGTACAGACTCCGGTCCACTCAAGGGGACAAATGG 629
|||
Db 1022 TCACATCAGAGGACCTCGTTTACCTCCGGGAGAAAGCTTCATGTTCCGACCTGGG 1081
|||
OY 630 TCACATCAGAGGACCTCGTTTACCTCCGGGAGAAAGCTTCATGTTCCGACCTGGG 689
|||
Db 1082 GGGCGGATGTTATCAACATGACACAGTTCACAGAGTGTTCTTGTAGAGAGCTGAA 1141
|||
OY 690 GGGCGGATGTTATCAACATGACACAGTTCACAGAGTGTTCTTGTAGAGAGCTGAA 749
|||
Db 1142 TTGTGTACGCAATATCGGCATGCGCACAGATTAATGACTGTGGAAGGACACAGAGAG 1201
|||
OY 750 TTGTGTACGCAATATCGGCATGCGCACAGATTAATGACTGTGGAAGGACACAGAGAG 809
|||
Db 1202 CAGT 1205
|||
OY 810 CAGT 813

RESULT 2
ID US-08-176-413-1 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Application US/08176413.
CC Sequence 1, Application US/08176413
CC Patent No. 5571510
CC GENERAL INFORMATION:
CC APPLICANT: No. 5571510ort, Tsutomu
CC APPLICANT: Carson, Dennis A.
CC TITLE OF INVENTION: METHOD FOR SELECTIVE METHIONINE
CC TITLE OF INVENTION: STARVATION OF MALIGNANT CELLS IN MAMMALS
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Spensley Horn Jubas & Lubitz
CC STREET: 1880 Century Park East, Suite 500
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90067
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/176,413
CC FILING DATE:
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Howells, Stacy L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: P22864
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 455-5100
CC TELEFAX: (619) 455-5100
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

```

```
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: methyladenosine phosphatase
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match      28.3%   Score 240; DB 5; Length 2763;
Best Local Similarity 99.28; Pred. No 9,33e-177;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cc
Db 962 AGGTTCTATAGAGACTGCTAAGAAGTAGACCTCGGTGCCACTCAAAAGGGGCACAATNG 1021
|
|
|
Qy 570 AGGTTCTTATAGAGACTGCTAAGAAGTAGACCTCGGTGCCACTCAAAAGGGGCACAATG 629
|
|
|
Db 1022 TCACATCAGAGGACCCTGTTTAAGCTCCGGCGAAGACCTTCATGTTCCGCACTTG 1081
|
|
|
Qy 630 TCACATCAGAGGACCCTGTTTAAGCTCCGGCGAAGACCTTCATGTTCCGCACTTG 689
|
|
|
Db 1082 GGGGGAGATTATCAACATGATGACACAGTTCAGAGAGTGTTTCTGCTAAGAGGCTGAA 1141
|
|
|
Qy 690 GGGGGAGATTATCAACATGATGACACAGTTCAGAGAGTGTTTCTGCTAAGAGGCTGAA 749
|
|
|
Db 1142 TTGTTACGAATATCGCATGGCACAGATATATGACTGCTGAAGAGCAGCAGAGAG 1201
|
|
|
Qy 750 TTGTTACGAATATCGCATGGCACAGATATATGACTGCTGAAGAGCAGCAGAGAG 809
|
|
|
Db 1202 CAGT 1205
|
|
|
Qy 810 CAGT 813
|
|
|
RESULT 3
ID PCT-US94-14920-1 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DE 01-JAN-1900
Sequence 1, Application PC/TUS9414920.
DE Sequence 1, Application PC/TUS9414920
CC GENERAL INFORMATION:
CC APPLICANT: THE REGENTS OF THE UNIVERSITY
CC APPLICANT: OF CALIFORNIA
CC TITLE OF INVENTION: METHOD FOR DETECTION OF
CC TITLE OF INVENTION: METHYLHIOADENOSINE PHOSPHATASE DEFICIENCY IN MANM
Cc ALIAN
CC TITLE OF INVENTION: CELLS
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Robbins, Berlinger & Carson
CC STREET: 201 N. Figueroa Street, 5th Floor
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90012
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/14920
Cc FILING DATE:
Cc CLASSIFICATION:
Cc ATTORNEY/AGENT INFORMATION:
Cc NAME: Berlinger, Robert
Cc REGISTRATION NUMBER: 20,121
Cc REFERENCE/DOCKET NUMBER: 5555-287
Cc TELECOMMUNICATION INFORMATION:
Cc TELEPHONE: 213-977-1001
Cc TELEFAX: 213-977-1003
Cc INFORMATION FOR SEQ ID NO. 1:
Cc SEQUENCE CHARACTERISTICS:
Cc LENGTH: 2763 base pairs
```


Tue Feb 11 11:27:33 1997

US-08-674-311-1-02.rni

Page 4

```

CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 215 base pairs
CC      type: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: unknown
CC      MOLECULE type: protein
CC      FEATURE:
CC      NAME/KEY: misc-feature
CC      LOCATION: 1..215
CC      OTHER INFORMATION: /standard_name="Deduced amino acid
CC      OTHER INFORMATION: sequence of PCp1 from bean."
CC
CC      Sequence 215 BP: 15 A; 8 C; 35 G; 26 T; 141 other;
CC

```

Query Match	3.3%	Score 28;	DB 5;	Length 215;
Best Local Similarity	16.9%;	Pred. No. 2.37e-05;		
Matches	21;	Conservative	49;	Mismatches 53;
			Indels	1;
			Gaps	1;

[illegible]

RESULT 6
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.

```

AC      01-JAN-1900      Application PC/TUS9511934.
DE      Sequence 100, Application PC/TUS9511934.
CC      Sequence 100, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC      APPLICANT: Cytogen Corporation
CC      TITLE OF INVENTION: Antigen Binding Peptides (Ab
CC      TITLE OF INVENTION: Peptide Libraries
CC      NUMBER OF SEQUENCES: 103
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESS: Pennie & Edmonds
CC      STREET: 1155 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: USA
CC      ZIP: 10036
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US95/11934
CC      FILING DATE: 20-SEP-1995
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Mistrick, S. Leslie
CC      REGISTRATION NUMBER: 18,872
CC      REFERENCE/DOCKET NUMBER: 1101-196-228
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (212) 790-9090
CC      TELEFAX: (212) 869-9741/8864
CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 100:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 74 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
CC      Sequence 74 bp; 6 A; 6 C; 1 G; 1 T; 60 other;

```

Query Match	2.98;	Score 25;	DB 11;	Length 74;
Best Local Similarity	9.08;	Pred. No. 1.71e-03;		
Matches	6;	Conservative	20;	Mismatches 41;
				Indels 0;
				Gaps 0.

D6 YNNVNNN 65

C6 CCGTGTCAAGAACTGATCATATAATGAACAATATGGCCGGCCTAATCTCTCCTCCAAGA 410

D6 ACCACAA 72
| | | | |
CP GGCACAA 403

RESULT	7
ID	PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.

AC *****
DT 01-JAN-1900
DE Sequence 99, Application PC/TU9511934.
CC Sequence 99, Application PC/TU9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Penite & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA

CC COMPUTER READABLE FORM:
CC
CC
CC MEDIUM TYPE: floppy disk
CC
CC COMPUTER: IBM PC compatible
CC
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC
CC CURRENT APPLICATION DATA:
CC
CC APPLICATION NUMBER: PCT/US95/11934
CC
CC FILING DATE: 20-SEP-1995
CC
CC CLASSIFICATION:
CC
CC ATTORNEY/AGENT INFORMATION:
CC
CC NAME: Mirock S. Leslie
CC
CC REGISTRATION NUMBER: 18,672
CC
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC
CC TELECOMMUNICATION INFORMATION:
CC
CC TELEPHONE: (212) 790-9090
CC
CC TELEFAX: (212) 869-9741/8864
CC
CC TELEX: 66141 PENNIE
CC
CC INFORMATION FOR SEQ ID NO: 99:
CC
CC SEQUENCE CHARACTERISTICS:
CC
CC LENGTH: 75 base pairs
CC
CC TYPE: nucleic acid
CC
CC STRANDEDNESS: single
CC
CC TOPOLOGY: linear
CC
CC MOLECULE TYPE: DNA (genomic)
CC
CC Sequence 75 Bp: 1 A; 1 C; 7 G; 5 T; 61 other;

Query Match	2.9%	Score 25;	DB 11;	Length 75;
Best Local Similarity	12.5%;	Pred. No. 1.71e-03;		
Matches	9;	Conservative 19;	Mismatches 44;	Indels 0;
			Gaps	0;

[illegible]

RESULT	8
ID	PCT-US95-11934-98 STANDARD; DNA; UNC; 81 BP.

CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 92:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 81 BP; 3 A; 5 C; 6 G; 4 T; 63 other;

Query Match 2.8%; Score 24; DB 11; Length 81;
Best Local Similarity 7.5%; Pred. No. 6.82e-03;
Matches 5; Conservative 20; Mismatches 42; Indels 0; Gaps 0;

Db 7 CTCGAGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNN 66
QY 412 CTGAGAGAGAGATTACCGCCGCGATATTGTCTATTGATCATGTGACAGGAC 471
Db 67 BNNBNNB 73
QY 472 CACTATG 478

RESULT 11
ID PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.
AC xxxxxx
DE Sequence 97, Application PC/TUS9511934.
CC Sequence 97, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELETYPE: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 97:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 82 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 82 BP; 1 A; 2 C; 10 G; 8 T; 61 other;

Query Match 2.8%; Score 24; DB 11; Length 82;
Best Local Similarity 7.5%; Pred. No. 6.82e-03;
Matches 5; Conservative 20; Mismatches 42; Indels 0; Gaps 0;

Db 7 CTCGAGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNN 66
QY 412 CTGAGAGAGAGATTACCGCCGCGATATTGTCTATTGATCATGTGACAGGAC 471

QY 412 CTGAGAGAGAGATTACCGCCGCGATATTGTCTATTGATCATGTGACAGGAC 471
Db 67 BNNBNNB 73
QY 472 CACTATG 478

RESULT 12
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DE Sequence 94, Application PC/TUS9511934.
CC Sequence 94, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELETYPE: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 94:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 74 BP; 3 A; 4 C; 3 G; 1 T; 63 other;

Query Match 2.7%; Score 23; DB 11; Length 74;
Best Local Similarity 4.7%; Pred. No. 2.64e-02;
Matches 3; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Db 3 GAGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNN 62
QY 415 GAGGAGAGAGATTACCGCCGCGATATTGTCTATTGATCATGTGACAGGACAC 474
Db 63 BNNB 66
QY 475 TATG 478

RESULT 13
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DE Sequence 100, Application PC/TUS9511934.
CC Sequence 100, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From

WIRELINE
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligent, Inc.

Msrch_un n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Feb 11 00:05:49 1997; MasPar time 337.15 Seconds
911.915 Million cell updates/sec

Tabular output not generated.

Title: >US-08-674-311-1
(122-970) from US08674311.seq (2 of 2)
Description: 849
Perfect Score: 122 ATGGCCCTGCGACACACAC.....CTGTTTATACCAAGACAT 970
N.A. Sequence: 122 TACCGAGACCGCTGCTGCTG.....GACAAATAATGCTCTGTA
Comp:

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 656168 seqs, 181066524 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

n-pending
1:PCR91 2:PCR92 3:PCR93 4:PCR94 5:PCR95 6:PCR96 7:U75
8:U76 9:U771 10:U772 11:U781 12:U782 13:U791 14:U792
15:U800 16:U801 17:U802 18:U803 19:U804 20:U805 21:U806
22:U807 23:U808 24:U809 25:U810 26:U811 27:U812 28:U813
29:U814 30:U815 31:U816 32:U817 33:U818 34:U819 35:U820
36:U821 37:U822 38:U823 39:U824 40:U825 41:U826 42:U827
43:U828 44:U829 45:U830 46:U831 47:U832 48:U833 49:U834
50:U835 51:U836 52:U837 53:U838 54:U839 55:U840 56:U841
57:U842 58:U843 59:U844 60:U845 61:U846 62:U847 63:U848
64:U849 65:U85 66:U86 67:U87 68:U88 69:U89 70:U90
71:NEW07 72:NEW08

Statistics: Mean 9.696; Variance 3.723; scale 2.604

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	849	100.0	2269	66	US-08-674-	Sequence 1, Applicatio	0.00e+00
2	849	100.0	2269	68	US-60-000-	Sequence 1, Applicatio	0.00e+00
3	240	28.3	2763	59	US-08-445-	Sequence 14, Applicati	2.29e-222
4	240	28.3	2763	32	US-08-176-	Sequence 1, Applicatio	2.29e-222
5	240	28.3	2763	59	US-08-443-	Sequence 14, Applicati	2.29e-222
6	240	28.3	2763	37	US-08-227-	Sequence 14, Applicati	2.29e-222
7	240	28.3	2763	37	US-08-227-	Sequence 14, Applicati	2.29e-222
8	173	20.4	4953	37	US-08-220-	Sequence 5097, Applic	4.96e-150
9	99	11.7	578	66	US-08-624-	Sequence 14, Applicati	3.37e-72
10	74	8.7	578	66	US-08-624-	Sequence 14, Applicati	5.66e-47

C	11	66	7.8	383	59	US-08-446-	Sequence 8, Applicatio	4.01e-39
C	12	66	7.8	383	46	US-08-311-	Sequence 8, Applicatio	4.01e-39
C	13	66	7.8	383	59	US-08-446-	Sequence 8, Applicatio	4.01e-39
C	14	66	7.8	383	59	US-08-446-	Sequence 8, Applicatio	4.01e-39
C	15	66	7.8	383	60	US-08-451-	Sequence 8, Applicatio	4.01e-39
C	16	66	7.8	383	59	US-08-446-	Sequence 8, Applicatio	4.01e-39
C	17	63	7.4	383	59	US-08-446-	Sequence 8, Applicatio	3.23e-36
C	18	63	7.4	383	46	US-08-446-	Sequence 8, Applicatio	3.23e-36
C	19	63	7.4	383	59	US-08-311-	Sequence 8, Applicatio	3.23e-36
C	20	63	7.4	383	59	US-08-451-	Sequence 8, Applicatio	3.23e-36
C	21	63	7.4	383	59	US-08-446-	Sequence 8, Applicatio	3.23e-36
C	22	63	7.4	383	59	US-08-446-	Sequence 8, Applicatio	3.23e-36
C	23	58	6.8	340	66	US-08-617-	Sequence 9, Applicatio	1.99e-31
C	24	57	6.7	339	66	US-08-617-	Sequence 9, Applicatio	1.77e-30
C	25	55	6.5	477	13	US-07-904-	Sequence 4, Applicatio	1.37e-28
C	26	54	6.4	477	13	US-07-904-	Sequence 4, Applicatio	1.19e-27
C	27	54	6.4	477	13	US-07-904-	Sequence 10, Applicati	1.19e-27
C	28	54	6.4	7218	38	US-08-232-	Sequence 14, Applicati	1.19e-27
C	29	54	6.4	7218	61	US-08-466-	Sequence 14, Applicati	1.19e-27
C	30	53	6.2	477	13	US-07-904-	Sequence 10, Applicati	1.03e-26
C	31	45	5.3	339	66	US-08-617-	Sequence 9, Applicati	2.26e-19
C	32	44	5.2	340	66	US-08-617-	Sequence 9, Applicati	1.78e-18
C	33	37	4.4	1645	8	US-07-697-	Sequence 68, Applicati	2.32e-12
C	34	36	4.2	137	55	US-08-404-	Sequence 32, Applicat	1.63e-11
C	35	35	4.1	1645	8	US-07-697-	Sequence 8, Applicati	1.63e-11
C	36	35	4.1	1645	8	US-07-697-	Sequence 8, Applicati	1.12e-10
C	37	34	4.0	965	53	US-07-904-	Sequence 6, Applicati	7.58e-10
C	38	34	4.0	965	53	US-08-388-	Sequence 58, Applicati	3.26e-08
C	39	32	3.8	222	62	US-07-904-	Sequence 8, Applicatio	3.26e-08
C	40	32	3.8	235	13	US-07-904-	Sequence 6, Applicatio	2.07e-07
C	41	31	3.7	235	13	US-08-388-	Sequence 22, Applicati	1.28e-06
C	42	30	3.5	965	53	US-08-292-	Sequence 133, Applicat	2.58e-04
C	43	27	3.2	66	65	US-08-515-	Sequence 5, Applicatio	2.58e-04
C	44	27	3.2	69	62	US-08-471-	Sequence 142, Applicat	2.58e-04
C	45	27	3.2					

ALIGNMENTS

RESULT 1
ID US-08-674-311-1 STANDARD; DNA; UNC; 2269 BP.
AC xxxxxx
DT 01-JAN-1990
DE Sequence 1, Application US/08674311.
CC Sequence 1, Application US/08674311
CC GENERAL INFORMATION:
CC APPLICANT: Olompe, Olufumlayo I.
CC TITLE OF INVENTION: METHYLTHIOADENOSINE PHOSPHORILASE
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE IN
CC TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF
CC NUMBER OF SEQUENCES: 15
CC PROLIFERATIVE DISORDERS
CC CORRESPONDENCE ADDRESS:
CC ADDRESSER: Arnold, White & Durkee
CC STREET: P.O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: United States of America
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/674.311
CC FILING DATE: Concurrently Herewith
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/000.831
CC FILING DATE: 02-JUL-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:

Tue Feb 11 11:27:35 1997

US-08-674-311-1-02.rtf

Page 2

CC NAME: Kitchell, Barbara S.
CC REGISTRATION NUMBER: 33,928
CC REFERENCE/DOCKET NUMBER: ARSB:509
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (512) 418-3000
CC TELEFAX: (512) 474-7577
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2269 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 122..970
CC Sequence 2269 BP; 725 A; 407 C; 490 G; 647 T; 0 other;

Query Match	100.0%;	Score 849;	DB 66;	Length 2269;
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;		
Matches	849;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0

Db	122	ATGGCCCTCTGGGACACACACACACACCGCGGTGAAATATGGAATATATGATGGAACAGGGCTG	181
Qy	122	ATGGCCCTCTGGGACACACACACACCGCGCGGTGAAATATGGAATATATGATGGAACAGGGCTG	181
Db	182	GATGATTCGAGAAATTTTGAAGGAGAAAGTGAATATATGATGATACCTCTTTGGGCAT	241
Qy	182	GATGATTCGAGAAATTTTGAAGGAGAAAGTGAATATATGATGATACCTCTTTGGGCAT	241
Db	242	CCATCTGATGCGCTTAATTTTGGGAAATATAAAATGTTGATGATCTCTCTGGCAAG	301
Qy	242	CCATCTGATGCGCTTAATTTTGGGAAATATAAAATGTTGATGATCTCTCTGGCAAG	301
Db	302	CATGAAGAGCAGCACACCATCATGCTCTCAAGAGTCACTACCAAGGCAAAATGTGGGCT	361
Qy	302	CATGAAGAGCAGCACACCATCATGCTCTCAAGAGTCACTACCAAGGCAAAATGTGGGCT	361
Db	362	TTGAGAGAAAGAGGCGCTGTACACATGTCATAGTACACACAGCTTGTGGCTCTTGAAGGAG	421
Qy	362	TTGAGAGAAAGAGGCGCTGTACACATGTCATAGTACACACAGCTTGTGGCTCTTGAAGGAG	421
Db	422	GAGATTGAGCCCGCGGCATATTGTCTATTATGATCAGTTCAATTGACAGACCATATAGAG	481
Qy	422	GAGATTGAGCCCGCGGCATATTGTCTATTATGATCAGTTCAATTGACAGACCATATAGAG	481
Db	482	CCTAGTGCCTTATATGAGAGTATATCTGTGCCAGAGAGATGCGCATATTCACATG	541
Qy	482	CCTAGTGCCTTATATGAGAGTATATCTGTGCCAGAGAGATGCGCATATTCACATG	541
Db	542	GCTAGAGCCTTTTGGCCCAAAACGAGAGAGGTTCTTTATAGACGTCCTTAAGAGGTAGGA	601
Qy	542	GCTAGAGCCTTTTGGCCCAAAACGAGAGAGGTTCTTTATAGACGTCCTTAAGAGGTAGGA	601
Db	602	CTCCGGGTGCCATCTCAAAGGGGACAAATGGTCAATTCAGGGACCTCGTTTATAGCTCCCG	661
Qy	602	CTCCGGGTGCCATCTCAAAGGGGACAAATGGTCAATTCAGGGACCTCGTTTATAGCTCCCG	661
Db	662	GCAGAAAGCTTCAATGTTTCGCGACCTGGGGGGCGGATGTTATCAATGACACAGTTTCCA	721
Qy	662	GCAGAAAGCTTCAATGTTTCGCGACCTGGGGGGCGGATGTTATCAAACTGACCAAGTTTCCA	721
Db	722	GAGGTGGTTTCTGCTAAGAGAGGCTGGAAATTTGTTACCAAGTATTCGCCATGGCGACAGAT	781
Qy	722	GAGGTGGTTTCTGCTAAGAGAGGCTGGAAATTTGTTACCAAGTATTCGCCATGGCGACAGAT	781
Db	782	TATGACTGCTGGAAGGAGCGACGAGAAAGATTTGGGTGGGCGGGGTCTTAAGAAGCCTTG	841
Qy	782	TATGACTGCTGGAAGGAGCGACGAGAAAGATTTGGGTGGGCGGGGTCTTAAGAAGCCTTG	841
Db	842	AAAGAAAACGCTAATTAAGCCAAAAAGCTTAACGTGCACACTACACTACTGATATAGAGTTC	901
Qy	842	AAAGAAAACGCTAATTAAGCCAAAAAGCTTAACGTGCACACTACTACTGATATAGAGTTC	901

Db	902	ACGAGTGGTCGAAACCCCTCATACCTGAAATAATAGCCAGTTTCGTGTATTA	961
Db	902	ACGAGTGGTCGAAACCCCTCATACCTGAAATAATAGCCAGTTTCGTGTATTA	961
Oy	902	ACGAGTGGTCGAAACCCCTCATACCTGAAATAATAGCCAGTTTCGTGTATTA	961
Db	962	CCAGAGCAT	970
Oy	962	CCAGAGCAT	970

```

ID      2
RESULT  US-60-000-831-1 STANDARD; DNA; UNC; 2269 BP.
AC      xxxxxx
DT      01-JAN-1900
DE      Sequence 1, Application US/60000831.
CC      Sequence 1, Application US/60000831
CC      GENERAL INFORMATION:
CC      APPLICANT: Olopade, Olufunmilayo I.
CC      APPLICANT: Bohlander, Stefan K.
CC      APPLICANT: Dreyling, Martin H.
CC      APPLICANT: Le Beau, Michelle M.
CC      APPLICANT: Stadler, Walter M.
CC      APPLICANT: Gursky, Susan
CC      APPLICANT: Espinosa, Ili, Rafael
CC      APPLICANT: Sveen, Lise W.
CC      APPLICANT: Hagos, Fitsum
CC      APPLICANT: Pomykala, Helen M.
CC      APPLICANT: Adeyanju, M.O.
CC      TITLE OF INVENTION: METHYLTHIOADENOSINE PHOSPHORYLASE
CC      TITLE OF INVENTION: METHODS AND COMPOSITIONS
CC      NUMBER OF SEQUENCES: 2
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Arnold, White & Durkee
CC      STREET: P.O. Box 4433
CC      CITY: Houston
CC      STATE: Texas
CC      COUNTRY: United States of America
CC      ZIP: 72210
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patentin Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/60/000,831
CC      FILING DATE: Concurrently Herewith
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Kitchell, Barbara S.
CC      REGISTRATION NUMBER: 33,928
CC      REFERENCE/DOCKET NUMBER: ARCD:156/KIT
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (512) 418-5000
CC      TELEFAX: (512) 474-7577
CC      INFORMATION FOR SEQ ID NO: 1:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 2269 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: other nucleic acid
CC      DESCRIPTION: /desc = "DNA"
CC      FEATURE:
CC      NAME/KEY: CDS
CC      LOCATION: 122..970
CC      Sequence 2269 BP; 725 A; 407 C; 490 G; 647 T; 0 other:

```

```
Query Match      score 849; DB 68; Length 2269;  
Best Local Similarity 100.0%; Pired. No. 0.00e+00;  
Matches    849; Conservative    0; Mismatches    0; Indels    0; Gaps    0.
```


Db	182	GATGATCCAGAAATTTTGAAGGAAGAACTGAAAAATATGTGATACATCCATTGGCAG	241
Qy	182	GATGATCCAGAAATTTTGAAGGAAGAACTGAAAAATATGTGATACATCCATTGGCAG	241
Db	242	CCATCTGATGCCCTTAATTTTGGGGAAGATTAATAATGTGATTGCATCCTCCTTCAAG	301
Qy	242	CCATCTGATGCCCTTAATTTTGGGGAAGATTAATAATGTGATTGCATCCTCCTTCAAG	301
Db	302	CATGGAAGGCAGACACCATCATGCTTCAAGGTCACACTACACAGCGCAACATCTGGGT	361
Qy	302	CATGGAAGGCAGACACCATCATGCTTCAAGGTCACACTACACAGCGCAACATCTGGGT	361
Db	362	TTGAAGGAAGAGGGCTGTACACATGTCATAGGACACACAGCTTGGGCTTGAAGGAG	421
Qy	362	TTGAAGGAAGAGGGCTGTACACATGTCATAGGACACACAGCTTGGGCTTGAAGGAG	421
Db	422	GAGATTCAGCCCCGGGATATTGTCTAATTATGATCAGTTTCATTGACAGGACCATATGAGA	481
Qy	422	GAGATTCAGCCCCGGGATATTGTCTAATTATGATCAGTTTCATTGACAGGACCATATGAGA	481
Db	482	CCTGATGCTCTTCTATGATGGAAGTATTTCTTGCCAGAGAGTGTCATATTCCATG	541
Qy	482	CCTGATGCTCTTCTATGATGGAAGTATTTCTTGCCAGAGAGTGTCATATTCCATG	541
Db	542	GCTGAGCCGTTTGGCCCAAAACGAGAGAGTTCCTTATAGACATGCTTAAAGCTAGGA	601
Qy	542	GCTGAGCCGTTTGGCCCAAAACGAGAGAGTTCCTTATAGACATGCTTAAAGCTAGGA	601
Db	602	CTCCGGTGCCACTCAAAGGGGACAAATGCTACAAATCGAGGGACCTCGTTTAGCTCCGG	661
Qy	602	CTCCGGTGCCACTCAAAGGGGACAAATGCTACAAATCGAGGGACCTCGTTTAGCTCCGG	661
Db	662	GCAGAAAGCTTCATTTTCCGCACCTGGGGGGGATGTTATTAACATGACCACTTCCA	721
Qy	662	GCAGAAAGCTTCATTTTCCGCACCTGGGGGGGATGTTATTAACATGACCACTTCCA	721
Db	722	GAGGTGTTTCTTGTCTAAGGAGGCTGGAATTTGTACGCAAGTACGCTGCGACAGAT	781
Qy	722	GAGGTGTTTCTTGTCTAAGGAGGCTGGAATTTGTACGCAAGTACGCTGCGACAGAT	781
Db	782	TATGACTCTGTGAAGGAGCAGCAGGAGCAAGTTCGTGGTGACCGGGCTTTAAAGACCTG	841
Qy	782	TATGACTCTGTGAAGGAGCAGCAGGAGCAAGTTCGTGGTGACCGGGCTTTAAAGACCTG	841
Db	842	AAAGAAACGCTAATAAAGCCAAAGCTTACGTCTACATACCATACCTCAGATAGGCTCC	901
Qy	842	AAAGAAACGCTAATAAAGCCAAAGCTTACGTCTACATACCATACCTCAGATAGGCTCC	901
Db	902	ACAGATAGGTCAAGAAACCTCCATTAACCTGAGAAATATGGCCAGATTTTCTGTTTATTA	961
Qy	902	ACAGATAGGTCAAGAAACCTCCATTAACCTGAGAAATATGGCCAGATTTTCTGTTTATTA	961
Db	962	CCAGAGACAT 970	
Qy	962	CCAGAGACAT 970	
RESULT 3			
ID	US-08-445-648-14	STANDARD: DNM: UNC; 2763 BP.	
AC	xxxxxx		
DT	01-JAN-1900		
DE	Sequence 14, Application US/08445648.		
CC	Sequence 14, Application US/08445648		
CC	GENERAL INFORMATION:		
CC	APPLICANT: CARSON, DENNIS A.		
CC	APPLICANT: NOBORI, TSUTOMU		
CC	TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR		
CC	TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESS		
CC	ON AND CANCER		
CC	TITLE OF INVENTION: TREATMENT		
CC	NUMBER OF SEQUENCES: 18		
CC	CORRESPONDENCE ADDRESS:		

```

CC ADDRESSSEE: Fish & Richardson P.C.
CC STREET: 4425 Executive Square, Suite 1400
CC CITY: La Jolla
CC STATE: California
CC COUNTRY: US
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/445,648
CC FILING DATE: 22-MAY-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/227,800
CC FILING DATE: 14-APR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: 07340/023001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/678-5070
CC TELEFAX: 619/678-5099
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC SEQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other:

Query Match 28.3%; Score 240; DB 59; Length 2763;
Best Local Similarity 99.2%; Pred.No.2,29e-222;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0

Db 962 AGGTTCTTATAGAGACTGCTAAGAAGCTAGAGACTCCGGTGCACCTCAAAGGGAGCAATGG 1021
|
|
|
Qy 570 AGGTTCTTATAGAGACTGCTAAGAAGCTAGAGACTCCGGTGCACCTCAAAGGGAGCAATGG 629
|
|
|
Db 1022 TCACATATGAGAGGACCTCGTTTACGCTCCGGGCGAAGACTTATATGTTCCGACCTGGG 1081
|
|
|
Qy 630 TCACATATGAGAGGACCTCGTTTACGCTCCGGGCGAAGACTTATATGTTCCGACCTGGG 689
|
|
|
Db 1082 GGGCGGATGTTATCAATGACACACAGTTCAGAGGCGTTCCTGCTAAGGAGGCTGGAA 1141
|
|
|
Qy 690 GGGCGGATGTTATCAATGACACACAGTTCAGAGGCGTTCCTGCTAAGGAGGCTGGAA 749
|
|
|
Db 1142 TTGTAGCAAGATATCGCATGGGACAGATATATGACTGCTGGAAGAGACGACAGGAAG 1201
|
|
|
Qy 750 TTGTAGCAAGATATCGCATGGGACAGATATATGACTGCTGGAAGAGAGACAGGAAG 809
|
|
|
Db 1202 CAGT 1205
|
|
|
Qy 810 CAGT 813

RESULT 4
ID US-08-176-855-1 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Application US/08176855.
CC Sequence 1, Application US/08176855.
CC GENERAL INFORMATION:
CC APPLICANT: Nobori, Tsutomu
CC APPLICANT: Carson, Dennis A.

```

Tue Feb 11 11:27:35 1997

US-08-674-311-1-02.rtf

Page 4

CC	APPLICANT:	Takabayashi, Kenji
CC	TITLE OF INVENTION:	METHOD FOR DETECTION OF
CC	TITLE OF INVENTION:	METHYLLITHIOADENOSINE PHOSPHATASE DEFICIENCY IN MANNA
CC	ALIAN	
CC	TITLE OF INVENTION:	CELLS
CC	NUMBER OF SEQUENCES:	5
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE:	Spensley Horn Jubas & Lubitz
CC	STREET:	1880 Century Park East, Suite 500
CC	CITY:	Los Angeles
CC	STATE:	California
CC	COUNTRY:	USA
CC	ZIP:	90067
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Floppy disk
CC	COMPUTER:	IBM PC compatible
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS
CC	SOFTWARE:	Patentin Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER:	US/08/176,855
CC	FILING DATE:	
CC	CLASSIFICATION:	435
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	Howells, Stacy L.
CC	REGISTRATION NUMBER:	34,842
CC	REFERENCE/DOCKET NUMBER:	PD0507
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	(619) 455-9100
CC	TELEFAX:	(619) 455-5110
CC	INFORMATION FOR SEQ ID NO:	1:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	2763 base pairs
CC	TYPE:	nucleic acid
CC	STRANDEDNESS:	single
CC	TOPOLOGY:	linear
CC	MOLECULE TYPE:	DNA (genomic)
CC	IMMEDIATE SOURCE:	
CC	CLONE:	methyladenosine phosphatase
CC	FEATURE:	
CC	NAME/KEY:	CDS
CC	LOCATION:	1..2763
CC	Sequence	2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ		
	Query Match	28.3%; Score 240; DB 32; Length 2763;
	Best Local Similarity	99.2%; Pred. No. 2,29e+22;
	Matches 242; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
Db	962 AGGTTCTTATAGAGACGCTTAGAAGCTAGACTCCGGTGCCACTCACAAGGGGCAATGC	1021
Oy	570 AGGTCTTAATAGAACACTGCTAGAAGCTAGACTCCGGTGCCACTCAAAGGGGCAAAATG	629
Db	1022 TCACATATGAGAGACGCTTAGAAGCTAGACTCCGGTGCCACTCATGTTCGCACTGGG	1081
Oy	630 TCACATATGAGAGACGCTTAGAAGCTAGACTCCGGTGCCACTCATGTTCGCACTGGG	689
Db	1082 GGGCGGAGTATATCAATATGACCACCACTTCCAAGGTGTTCTTGCTAAGAGGCTGA	1141
Oy	690 GGGCGGAGTATATCAATATGACCACCACTTCCAAGGTGTTCTTGCTAAGAGGCTGA	749
Db	1142 TTGTGTACGCAAGTATGCCATGGGCGCACAGATTATGACTGCTGGAAGGACACAGAG	1201
Oy	750 TTGTGTACGCAAGTATGCCATGGGCGCACAGATTATGACTGCTGGAAGGACACAGAG	809
Db	1202 CAGT 1205	
Oy	810 CAGT 813	
RESULT	5	
ID	US-08-443-880-14 STANDARD; DNA; UNC; 2763 BP.	
AC	xxxxxx	
DT	01-JAN-1900	
DE	Sequence 14, Application US/08443880.	

CC	Sequence 14, Application US/08443880
CC	GENERAL INFORMATION:
CC	APPLICANT: CARSON, DENNIS A.
CC	APPLICANT: NOBORI, TSUTOMU
CC	TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR
CC	TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESS
CC	ON AND CANCER
CC	TITLE OF INVENTION: TREATMENT
CC	NUMBER OF SPOUNCES: 18
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Fish & Richardson P.C.
CC	STREET: 4225 Executive Square, Suite 1400
CC	CITY: La Jolla
CC	STATE: California
CC	COUNTRY: US
CC	ZIP: 92037
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/443,880
CC	FILING DATE: 22-MAY-1995
CC	CLASSIFICATION: 530
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: 08/227,800
CC	FILING DATE: 14-APR-1994
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: HOWELLS, STACY L.
CC	REGISTRATION NUMBER: 34,842
CC	REFERENCE/DOCKET NUMBER: 07340/023001
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: 619/678-5070
CC	TELEFAX: 619/678-5089
CC	INFORMATION FOR SEQ ID NO: 14:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 2763 base pairs
CC	TYPE: nucleic acid
CC	STRANDEDNESS: single
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: DNA (genomic)
CC	IMMEDIATE SOURCE:
CC	CLONE: Methylothioadenosine phosphorylase (genomic)
CC	FEATURE:
CC	NAME/KEY: CDS
CC	LOCATION: 1..2763
SO	Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
	Query Match 28.3%; Score 240; DB 59; Length 2763;
	Best Local Similarity 99.2%; Pred. No. 2,29e+22;
	Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db	962 AGGTTCTTATAGACAGCGTAAGAAGCTAAGAGATCGGGTCGCATCAAGAGGGACAATG 1021
OY	570 AGGTTCTTATAGAGCATGCTTAGAGAGCTTAGAGACTCCGGTGCCACTCAAGGGACAATG 629
Db	1022 TCACATCGAGGGAGCCTGTGTTTTAGCTCCCGGCCAANAAGCTTCACTGTTCCGCACTGG 1081
OY	630 TCACATCGAGGGAGCCTGTGTTTTAGCTCCCGGCCAANAAGCTTCACTGTTCCGCACTGG 689
Db	1082 GGCGGAGATGTTATCAACATGACCACAGATTCCAGAGGTGGTCTTGCTAAGAGAGCTGAA 1141
OY	690 GGCGGAGATGTTATCAACATGACCACAGATTCCAGAGGTGGTCTTGCTAAGAGAGCTGAA 749
Db	1142 TTGTGTAGCAAGTATCGCATCGGACAGATTATGACTGCTGGAAGGACGACGAGAG 1201
OY	750 TTGTGTAGCAAGTATCGCATCGGACAGATTATGACTGCTGGAAGGACGACGAGAG 809
Db	1202 CAGT 1205
OY	810 CAGT 813

RESULT 6
ID US-08-227-800A-14 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08227800A.
CC Sequence 14, Application US/08227800A
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.
CC APPLICANT: NOBORI, TSUTOMU
CC TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR
CC TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESS
ON AND CANCER
CC TITLE OF INVENTION: TREATMENT
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 4225 Executive Square, Suite 1400
CC CITY: La Jolla
CC STATE: California
CC COUNTRY: US
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/227,800A
CC FILING DATE: 14-APR-1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: 07340/023001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/678-5070
CC TELEFAX: 619/678-5099
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other:
Query Match 28.3%; Score 240; DB 37; Length 2763;
Best Local Similarity 99.2%; Pred. No. 2,29e-222;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 962 AGTTCCTTATAGAGCTGTAGAGAGCTAGAGCTCCGGTGCACCTCAAGGGAGCAATGG 1021
Oy 570 AGTTCCTTATAGAGCTGTAGAGAGCTAGAGCTCCGGTGCACCTCAAGGGAGCAATGG 629
Db 1022 TCACATGAGGAGGACCTGTTTACCTCCGGGAGAAAGCTTCATGTTCCGACCTGGG 1081
Oy 630 TCACATGAGGAGGACCTGTTTACCTCCGGGAGAAAGCTTCATGTTCCGACCTGGG 689
Db 1082 GGGCGAGTGTATCAACATGACACAGTTCAGAGGTGTTCTTGTGAAGAGGCTGGA 1141
Oy 690 GGGCGAGTGTATCAACATGACACAGTTCAGAGGTGTTCTTGTGAAGAGGCTGGA 749
Db 1142 TTGTAGCAGAGTATCGGCATGGGACAGATTATGACTGCTGGAAGAGCAGCAGGAAG 1201
Oy 750 TTGTAGCAGAGTATCGGCATGGGACAGATTATGACTGCTGGAAGAGCAGCAGGAAG 809
Db 1202 CAGT 1205

Oy 810 CAGT 813
RESULT 7
ID US-08-227-800-14 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08227800.
CC Sequence 14, Application US/08227800
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.
CC APPLICANT: NOBORI, TSUTOMU
CC TITLE OF INVENTION: METHODS FOR TREATMENT AND DETECTION OF
CC TITLE OF INVENTION: SUSCEPTIBILITY TO, AND PROGRESSION OF, CANCER
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
CC STREET: 1880 Century Park East, Suite 500
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: US
CC ZIP: 90067
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/227,800
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: PD3459
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/455-5100
CC TELEFAX: 619/455-5110
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other:
Query Match 28.3%; Score 240; DB 37; Length 2763;
Best Local Similarity 99.2%; Pred. No. 2,29e-222;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 962 AGTTCCTTATAGAGCTGTAGAGAGCTAGAGCTCCGGTGCACCTCAAGGGAGCAATGG 1021
Oy 570 AGTTCCTTATAGAGCTGTAGAGAGCTAGAGCTCCGGTGCACCTCAAGGGAGCAATGG 629
Db 1022 TCACATGAGGAGGACCTGTTTACCTCCGGGAGAAAGCTTCATGTTCCGACCTGGG 1081
Oy 630 TCACATGAGGAGGACCTGTTTACCTCCGGGAGAAAGCTTCATGTTCCGACCTGGG 689
Db 1082 GGGCGAGTGTATCAACATGACACAGTTCAGAGGTGTTCTTGTGAAGAGGCTGGA 1141
Oy 690 GGGCGAGTGTATCAACATGACACAGTTCAGAGGTGTTCTTGTGAAGAGGCTGGA 749
Db 1142 TTGTAGCAGAGTATCGGCATGGGACAGATTATGACTGCTGGAAGAGCAGCAGGAAG 1201
Oy 750 TTGTAGCAGAGTATCGGCATGGGACAGATTATGACTGCTGGAAGAGCAGCAGGAAG 809

QY 588 CTAAGAA-CTAGAGATCCGGTCC-CACCTAAGGGGACATGTCACATCAGGAGACC 645
Db 301 GATSVKNGGTVDSGRCAVATKTRKSTRYHSSCVKSGNGSVKTRKSDYDHKSSDSDS 360
QY 646 TCGTTTACCTCCCGGAGAAAGCTTCAATGTTCCACCTGGGGCGGATGTATCA 705
Db 361 YGGCAATTTTTHGTHGKTSKTRKYDKTDSGK-TKTSHEMKVTSACGSSST 419
QY 706 CATGACACAGTTCACAGAGTGTCTTCTGTAAGGAGGCTGGAATTTGTACCAAGTAT 765
Db 420 NNYMGVGHVHTSDPAMSYDSCGSGTSGKSTANRGRTSRSDKNAAGTDTMTNG 479
QY 766 CCCCATGGGACAGATTAAGACTGCTGGAAGAGCAGACAGAGAGAGTTCGGTGACCG 825
Db 480 SSSKARKDSKMRKCGSSVTKADRGGKVAARSTCSSSAGRAAGDKSGKRTVSSAV-AK 538
QY 826 GGTCTTAAAGACCTGAAGAAACGCTATATAAGCCAAAGCTTACTGTACATCAT 885
Db 539 RKSRSNTTRKVSVAHHSKGRD 563
QY 886 ACCTCAGATAGGCTCCACAGAAAG 910

RESULT 10
ID US-08-624-190-14 STANDARD; DNA; UNC; 578 BP.

AC XXXXX
DT 01-JAN-1900
DE Sequence 14, Application US/08624190.
CC Sequence 14, Application US/08624190.
CC GENERAL INFORMATION:
CC APPLICANT: KAVANAUGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
CC TITLE OF INVENTION: 5-PHOSPHATASE
CC NUMBER OF SEQUENCES: 38
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHIRON CORPORATION
CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/624,190
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.003
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-655-3542
CC INFORMATION FOR SEQ. ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC Sequence 578 BP; 37 A; 11 C; 62 G; 65 T; 403 other;

Query Match 8.7%; Score 74; DB 66; Length 578;
Best Local Similarity 11.7%; Pred. No. 5,66e-47;
Matches 48; Conservative 207; Mismatches 147; Indels 9; Gaps 9;

Db 3 SARVTSRSDSGRRVANNMWSKSTSSDPKYAHGSHRSVYKASGKMKVDVSGKKSD 62
CP 679 GAACATGAAGCTTTTCGCCCGGAGCTAAACAGAGCTCCCTGATG-TGACCATTTGCC 621
Db 63 GSDPKSHKRRKKNKYTKRKRYVADSKRKGCKN-KHSDMTGTMWNAKRTSMKSGRT 121
CP 620 CCTTTAGGAGGACCGGAGTCTTCTTGTAGCAGTCTCTATA-AGAACCCTCTCGTT 562
Db 122 RDDSADYHDYVGTDSKMKHSTSVTKVAHTMNRVYAKHNRSHCTDNVKTGANTGNKAVG 181
CP 561 TTGGGCAAAAGAGCTC-AGCCATGGAATATGACACACTCTGGCACAAGATGAGT 503
Db 182 VSMNGSSGVNSHTSGSKRRNNYMNRAKDKSNTHTHNGDVRPDYATKRYADSHDR 241
CP 502 TCCTA-TCAATAGAAGGACTAGGCTCATATGCTGCTGCAATGAACATGAT-CAATTAAG 445
Db 242 RKVHTATYRTDHYVTKATGMKYNMCDPVMKSYVWVSYSGSTDMTSDHVAATG 301
CP 444 ACAATATCCCGGCGGCTGAATCTCTCTCTCAAGAGCCACAAGCTGTGG-TCATCATGAC 386
Db 302 VTSVSKNGGTVDSGRCAVATKTRKSTRYHSSCVKSGNGSVKTRKSDYDHKSSDSDSY 361
CP 385 ATGTGTACAGCCCTCTCTCTCAAGCCAGATGTT-CGCTGTGATGTGACCTTTGAAG 327
Db 362 GGCARATTTTTHGTHGKTSKTRKYDKTDSGKTSKTSKTHDKMWTSR 412
CP 326 GCATGATGTTGTGCTGCTTCATGCTTGCATGCAAGAG-GATGCAATCAACA 277

RESULT 11
ID US-08-446-939-8 STANDARD; DNA; UNC; 383 BP.

AC XXXXX
DT 01-JAN-1900
DE Sequence 8, Application US/08446939.
CC Sequence 8, Application US/08446939.
CC GENERAL INFORMATION:
CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF- β Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,939
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI91-09FX
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240

CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 383 bp; 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match 7.88; Score 66; DB 59; Length 383;
Best Local Similarity 11.98; Pred. No. 4.01e-39;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;

Db 1 MTSSGCHGSGAGMHWYTRASTHVSVNNDVTDNNGAVKCKDVRSTCDKSCMSNCS 60
CP ATCCGCCCCCAGGTGGGACATGAAGCTTCTGCCGGGAGTAAACAGAGTCCCTC 638
Db 61 TSCRYCAVWRKNDNTVCHDKYHDDAASKMKKKKGTMCSCSDCNDNSYNTSNDVYT 120
CP GATGTGACCATTTGCTCCCTTTAGTGACGACCGAGTCTTACTCTTAGCAGTCTC-TA 579
Db 121 SGVAVYCYRVNRKSSYTWGTRKMSHCADDRSDSTCANNNHNDTYGKRAYYKAKN 180
CP 578 T-AMGACCTCTCGTTGGGCAAAAGGCTCAGCATTTGAATA-TGGCACACTCC 521
Db 181 TSTVAVYDYHASKRDKSDNKHNTARKTGKYWTAAKGNTRHVSMDRVNGSSARSH 240
CP 520 TCTGGCACAAATAGATTCATCATGAGAAGACTAGAGTCTCATAGTGTCTGTCAAT 461
Db 241 HSDHTGCRKRVHRDKSSNVKNDTCCDGRGYSVDANSVGTARYAVSMNNASTD 300
CP 460 GAACGTGATCAATAATGACAATATCGCGGCTGAATCTCTCCCT-CAAGGACCCACAG 402
Db 301 VYSMAVMTSRCAVGVKDYGSKRVADVSMKDNVRDGRNSNMHNGCTTCMDHART 360
CP 401 CTGTGTCATATGACATGTGTACAGCCCTCTTCAAGCCACAGATGTTCGCTGT 342
Db 361 A 361
CP 341 A 341

RESULT 12
ID US-08-311-703-8 STANDARD; DNA; UNC; 383 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 8, Application US/08311703.
CC Sequence 8, Application US/08311703
CC GENERAL INFORMATION:
CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF-beta Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P. C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC FILING DATE: US/08/311.703
CC CLASSIFICATION: 530

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WH191-09
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-9540
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 383 bp; 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match 7.88; Score 66; DB 46; Length 383;
Best Local Similarity 11.98; Pred. No. 4.01e-39;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;

Db 1 MTSSGCHGSGAGMHWYTRASTHVSVNNDVTDNNGAVKCKDVRSTCDKSCMSNCS 60
CP ATCCGCCCCCAGGTGGGACATGAAGCTTCTGCCGGGAGTAAACAGAGTCCCTC 638
Db 61 TSCRYCAVWRKNDNTVCHDKYHDDAASKMKKKKGTMCSCSDCNDNSYNTSNDVYT 120
CP GATGTGACCATTTGCTCCCTTTAGTGACGACCGAGTCTTACTCTTAGCAGTCTC-TA 579
Db 121 SGVAVYCYRVNRKSSYTWGTRKMSHCADDRSDSTCANNNHNDTYGKRAYYKAKN 180
CP 578 T-AMGACCTCTCGTTGGGCAAAAGGCTCAGCATTTGAATA-TGGCACACTCC 521
Db 181 TSTVAVYDYHASKRDKSDNKHNTARKTGKYWTAAKGNTRHVSMDRVNGSSARSH 240
CP 520 TCTGGCACAAATAGATTCATCATGAGAAGACTAGAGTCTCATAGTGTCTGTCAAT 461
Db 241 HSDHTGCRKRVHRDKSSNVKNDTCCDGRGYSVDANSVGTARYAVSMNNASTD 300
CP 460 GAACGTGATCAATAATGACAATATCGCGGCTGAATCTCTCCCT-CAAGGACCCACAG 402
Db 301 VYSMAVMTSRCAVGVKDYGSKRVADVSMKDNVRDGRNSNMHNGCTTCMDHART 360
CP 401 CTGTGTCATATGACATGTGTACAGCCCTCTTCAAGCCACAGATGTTCGCTGT 342
Db 361 A 361
CP 341 A 341

RESULT 13
ID US-08-446-936-8 STANDARD; DNA; UNC; 383 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 8, Application US/08446936.
CC Sequence 8, Application US/08446936
CC GENERAL INFORMATION:
CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF-Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P. C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,936
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI91-09FW
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 383 BP: 27 A; 25 C; 26 G; 34 T; 271 other:
SO
Query Match 7.8%; Score 66; DB 59; Length 383;
Best Local Similarity 11.9%; Pred. No. 4.01e-39;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;
Db 1 MTSSGVCAGAGWVHTRASTHKSVMNDVTDNNGAVKCDVRSCTDNKSCMSNCS 60
Cp ATCCGCCCCCAGAGCGGAGATGAGCTTCTCCCGGAGCTAAACGAGCTCCCTC 638
Db 61 TSCKVCVAVMRKNDTVCHDKYHDDAASKCMKKKGTWCSSSDCNDNSYNTSNDVYTG 120
Cp 637 GATGTGACATATGACAAATATGCGCGGCTGAATCTCTCCCT-CAAGGAGCCACAAG 402
Db 121 SGVAVYCYRVRNRSSTGTGTRKMSHCADRSSTTCANNHNTDYGKRAYKAKKN 180
Cp 578 T-AAGAACCTCTCTGCTTTGGGCAAAACGGCTCAGCATTTGAATA-TGGCACACTCC 521
Db 181 TSTVAVKTDHYASWDRKSDNKHNTARKTGKWTAAHAKGNTTRHVSMDRVNGSSARSGH 240
Cp 520 TCTGGCACAAGAAATGACTTCATCATAGAGAGCTGAGGTCTCATAGTGGCTCTCATAT 461
Db 241 HSDTCGRKMYHRDKSSVNVKNDTCCCGDSRGYSYVDANSGVGTARYMAVSRMNAKST 300
Cp 460 GAACGTGATCAATATGACAAATATGCGCGGCTGAATCTCTCCCT-CAAGGAGCCACAAG 402
Db 301 VYSMAVMTSRNAVGVNDYGSKVDDVYSMKNDVDRGTRNNSNMHGVCTTCMDHDKRT 360
Cp 401 CTGTGGTACATATGACAAATATGCGCGGCTGAATCTCTCCCT-CAAGGAGCCACAAG 402
Db 361 A 361
Cp 341 A 341

CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF- Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,937
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI91-09FW
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 383 BP: 27 A; 25 C; 26 G; 34 T; 271 other:
SO
Query Match 7.8%; Score 66; DB 59; Length 383;
Best Local Similarity 11.9%; Pred. No. 4.01e-39;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;
Db 1 MTSSGVCAGAGWVHTRASTHKSVMNDVTDNNGAVKCDVRSCTDNKSCMSNCS 60
Cp ATCCGCCCCCAGAGCGGAGATGAGCTTCTCCCGGAGCTAAACGAGCTCCCTC 638
Db 61 TSCKVCVAVMRKNDTVCHDKYHDDAASKCMKKKGTWCSSSDCNDNSYNTSNDVYTG 120
Cp 637 GATGTGACATATGACAAATATGCGCGGCTGAATCTCTCCCT-CAAGGAGCCACAAG 402
Db 121 SGVAVYCYRVRNRSSTGTGTRKMSHCADRSSTTCANNHNTDYGKRAYKAKKN 180
Cp 578 T-AAGAACCTCTCTGCTTTGGGCAAAACGGCTCAGCATTTGAATA-TGGCACACTCC 521
Db 181 TSTVAVKTDHYASWDRKSDNKHNTARKTGKWTAAHAKGNTTRHVSMDRVNGSSARSGH 240
Cp 520 TCTGGCACAAGAAATGACTTCATCATAGAGAGCTGAGGTCTCATAGTGGCTCTCATAT 461
Db 241 HSDTCGRKMYHRDKSSVNVKNDTCCCGDSRGYSYVDANSGVGTARYMAVSRMNAKST 300
Cp 460 GAACGTGATCAATATGACAAATATGCGCGGCTGAATCTCTCCCT-CAAGGAGCCACAAG 402
Db 301 VYSMAVMTSRNAVGVNDYGSKVDDVYSMKNDVDRGTRNNSNMHGVCTTCMDHDKRT 360
Cp 401 CTGTGGTACATATGACAAATATGCGCGGCTGAATCTCTCCCT-CAAGGAGCCACAAG 402
Db 361 A 361
Cp 341 A 341

(TM)

Distribution rights by Intelligenetics, Inc

n.a. database search, using Smith-Waterman algorithm

Mon Feb 10 23:14:27 1997; MasPar time 85.57 Seconds

c generated.

entire seq
674311.seq (1 c

1 GAATTCGCTCCGCACTGCT.....AAAAGGCGGATTC 2269

CTTAAGCGAGGCCGTGACGA.....TTTTTTTTTTCGCCCTTAAG

TABLE default

Gap 6

50278 seqs, 13829960 bases x 2

Minimum Match 0%

n-issued

Mean 9.387; Variance 5.460; scale 1.719

Under the number of results predicted by chance to have a better than or equal to the score of the result being printed, divided by analysis of the total score distribution.

SUMMARIES

Entry	Length	DB	ID	Description	Pred. No.
20.5	2763	10	PCr-US94-1	Sequence 1, Applicatio	0.00e+00
20.5	2763	5	US-08-176-	Sequence 1, Applicatio	0.00e+00
20.5	2763	10	PCr-US94-1	Sequence 1, Applicatio	0.00e+00
1.5	215	5	US-08-238-	Sequence 5, Applicatio	5.79e-06
1.4	215	5	US-08-238-	Sequence 5, Applicatio	7.16e-05
1.1	74	11	PCr-US95-1	Sequence 10, Applicat	2.80e-01
1.1	74	11	PCr-US95-1	Sequence 10, Applicat	9.13e-02
1.1	75	11	PCr-US95-1	Sequence 99, Applicat	9.13e-02
1.1	81	11	PCr-US95-1	Sequence 99, Applicat	2.80e-01
1.1	81	11	PCr-US95-1	Sequence 92, Applicat	9.13e-02
1.1	82	11	PCr-US95-1	Sequence 97, Applicat	2.80e-01
1.1	82	11	PCr-US95-1	Sequence 97, Applicat	9.13e-02
1.1	3214	5	US-08-484-	Sequence 17, Applicat	2.80e-01
1.0	26	12	PCr-US96-0	Sequence 269, Applicat	4.45e+00
1.0	74	11	PCr-US95-1	Sequence 94, Applicat	8.40e-01
1.0	75	11	PCr-US95-1	Sequence 94, Applicat	2.45e+00
1.0	81	11	PCr-US95-1	Sequence 92, Applicat	2.45e+00

15	21	0.9	3291	10	PCT-US94-0	Sequence 1,	Applicatio	6.97e+00
44	21	0.9	3138	3	US-07-867-	Sequence 4,	Applicatio	6.97e+00
43	21	0.9	2511	5	US-08-353-	Sequence 1,	Applicatio	6.97e+00
24	22	1.0	1952	5	US-08-333-	Sequence 1,	Applicatio	2.45e+00
23	22	1.0	1841	11	PCT-US93-0	Sequence 1,	Applicatio	2.45e+00
25	22	1.0	2148	8	PCT-US92-0	Sequence 12,	Applicati	2.45e+00
26	22	1.0	2504	5	US-08-484-	Sequence 15,	Applicati	2.45e+00
27	22	1.0	2653	5	US-08-335-	Sequence 1,	Applicati	2.45e+00
28	23	1.0	2913	5	US-08-243-	Sequence 7,	Applicatio	8.40e-01
29	23	1.0	2923	5	US-08-243-	Sequence 6,	Applicatio	8.40e-01
30	22	1.0	3075	10	PCT-US94-0	Sequence 1,	Applicatio	2.45e+00
31	22	1.0	3034	9	PCT-US93-1	Sequence 3,	Applicatio	2.45e+00
32	23	1.0	3183	5	US-08-243-	Sequence 8,	Applicatio	8.40e-01
33	22	1.0	3387	11	PCT-US95-0	Sequence 5,	Applicatio	8.40e-01
34	22	1.0	3933	11	PCT-US95-0	Sequence 1,	Applicatio	2.45e+00
35	22	1.0	3933	5	US-08-139-	Sequence 1,	Applicatio	2.45e+00
36	23	1.0	8342	10	PCT-US94-0	Sequence 63,	Applicati	8.40e-01
37	23	1.0	8392	9	PCT-US93-0	Sequence 6,	Applicatio	8.40e-01
38	23	1.0	8392	4	US-08-080-	Sequence 6,	Applicatio	8.40e-01
39	22	1.0	1557	8	PCT-US92-0	Sequence 1,	Applicatio	2.45e+00
40	21	0.9	105	4	US-07-885-	Sequence 13,	Applicati	6.97e+00
41	21	0.9	2258	3	US-07-770-	Sequence 1,	Applicatio	6.97e+00
42	21	0.9	2504	5	US-08-484-	Sequence 15,	Applicati	6.97e+00
43	21	0.9	2511	5	US-08-353-	Sequence 1,	Applicatio	6.97e+00
44	21	0.9	3138	3	US-07-867-	Sequence 4,	Applicatio	6.97e+00
45	21	0.9	3291	10	PCT-US94-0	Sequence 1,	Applicatio	6.97e+00

ALIGNMENTS

ID RESULT 1
 PC/US94-14919-1 STANDARD; DNA; UNC; 2763 BP.
 AC xxxxxx
 DT 01-JAN-1900
 DE Sequence 1, Application PC/TUS9414919.
 CC Sequence 1, Application PC/TUS9414919
 CC GENERAL INFORMATION:
 CC APPLICANT: THE REGENTS OF THE UNIVERSITY
 CC APPLICANT: OF CALIFORNIA
 CC TITLE OF INVENTION: METHOD FOR SELECTIVE METHYLATION
 CC TITLE OF INVENTION: STARVATION OF MALIGNANT CELLS IN MAMMALS
 CC NUMBER OF SEQUENCES: 5
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Robbins, Berliner & Carson
 CC STREET: 201 N. Figueroa Street, 5th Floor
 CC CITY: Los Angeles
 CC STATE: California
 CC COUNTRY: USA
 CC ZIP: 90012
 CC
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PC/US94/14919
 CC FILING DATE:
 CC CLASSIFICATION:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Berliner, Robert
 CC REGISTRATION NUMBER: 20,121
 CC REFERENCE/DOCKET NUMBER: 5555-286
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 213-977-1001
 CC TELEFAX: 213-977-1003
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 2763 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: DNA (genomic)
 CC IMMEDIATE SOURCE:

Tue Feb 11 11:27:26 1997

US-08-674-311-1-01.rni

Page 2

CC	CLONE:	methylenediosine phosphatase
CC	FEATURE:	
CC	NAME/KEY:	CDS
CC	LOCATION:	1..2763
CC	Sequence	2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
QY	Query Match	20.5%; Score 466; DB 10; Length 2763;
QY	Best Local Similarity	99.0%; Pred. NO. 0.00e+00;
QY	Matches 481; Conservative	0; Mismatches 1; Indels 4; Gaps
Db	2270	AGATATATGCCAGTTCTTGTTTATTACCAAGACATTAAAGTAGCATGCTGCCAGG 2329
QY	933	AGAAATATAGGCCAGTTCTTGTTTATTACCAAGACATTAAAGTAGCATGCTGCCAGG 992
Db	2330	AGAAAGAAACATCTTAATTCAGTCATTTTGGGAATTCCTGCTTAACCTGAAAAAAT 2389
QY	993	AGAAAGAAACATCTTAATTCAGTCATTTTGGGAATTCCTGCTTAACCTGAAAAAAT 10522
Db	2390	ATGGGAAGACATGACACTTTCATGCGCTTGCCTATCAAGAAGTATGTTGTAGAAAGAC 2449
QY	1053	ATGGGAAGACATGACACTTTCATGCGCTTGCCTATCAAGAAGTATGTTGTAGAAAGAC 11122
Db	2450	AAGCATTGTGTGAT-AGAGACTCCTCAATATTTAGCAACTCAAAATACAGAGAA 25088
QY	1113	AAGCATTGTGTGATAGAGACTCCTCAATATTTAGCAACTCAAAATACAGAGAA 11727
Db	2509	AAGCAATGACTAGTA-CATGTGGGAAAAAATATTACATTTTAAGGGGAAAAA - 25555
QY	1173	AAGCAATGACTAGTAACATGTGGGAAAAAATATTACATTTTAAGGGGAAAAA 12322
Db	2566	CCCCACATTCTCTTCCCCCATTTAAATTTTGCACATTAAGGGTGGAGGTAATCTC 26259
QY	1233	CCCCACATTCTCTTCCCCCATTTAAATTTTGCACATTAAGGGTGGAGGTAATCTC 12922
Db	2626	TACTTCTTACGCGCAAGATGTGAGGAAGAAATGGAGCTCTTGGTTTATTATGA 26858
QY	1293	TACTTCTTACGCGCAAGATGTGAGGAAGAAATGGAGCTCTTGGTTTATTATGA 13523
Db	2686	TGCGACATCTTAATTTGTCACGTAATTTCTGAGGGCAATTTGTTAAATGCATCAAAAC 27455
QY	1353	TGCGACATCTTAATTTGTCACGTAATTTCTGAGGGCAATTTGTTAAATGCATCAAAAC 14122
Db	2746	TTAAAAATACGAGCT 2761
QY	1413	TTAAAAATACGAGCT 1428
RESULT	2	
ID	US-08-176-413-1	STANDARD; DNA-ENG; 2763 BP.
AC	xxxxxx	
DT	01-JAN-1900	
DE	Sequence 1, Application US/08176413.	
CC	Sequence 1, Application US/08176413	
CC	Patent No. 5571510	
CC	GENERAL INFORMATION:	
CC	APPLICANT: NO. 5571510ori1, Tsutomu	
CC	APPLICANT: Carlson, Dennis A.	
CC	TITLE OF INVENTION: METHOD FOR SELECTIVE METHIONINE	
CC	NUMBER OF SEQUENCES: 5	
CC	STARAVATION OF MALIGNANT CELLS IN MAMMALS	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: Spensley Horn Jubas & Lubitz	
CC	STREET: 1880 Century Park East, Suite 500	
CC	CITY: Los Angeles	
CC	STATE: California	
CC	COUNTRY: USA	
CC	ZIP: 90067	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: Patentin Release #1.0, Version #1.25	
CC	CURRENT APPLICATION DATA:	

CC	APPLICATION NUMBER: US/08/176,413
CC	FILING DATE:
CC	CLASSIFICATION: 424
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Howells, Stacy L.
CC	REGISTRATION NUMBER: 34,842
CC	REFERENCE/DOCKET NUMBER: PD2664
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (619) 455-5100
CC	TELEFAX: (619) 455-5110
CC	INFORMATION FOR SEQ ID NO: 1:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 2763 base pairs
CC	TYPE: nucleic acid
CC	STRANDEDNESS: single
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: DNA (genomic)
CC	IMMEDIATE SOURCE:
CC	CLONE: methyladenosine phosphatase
CC	FEATURE:
CC	NAME/KEY: CDS
CC	LOCATION: 1..2763
CC	Sequence 2763 bp; 760 A; 505 C; 598 G; 803 T; 77 other:
CC	
CC	Query Match 20.5%; Score 466; DB 5; Length 2763;
CC	Best Local Similarity 99.0%; Pred. No. 0.00e+00;
CC	Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3.
Db	2270 AGAATATGGCCCAAGTTCTTGTTTATTATTCACAGACATTAAGTAGCATGCTGCCAGG 2329
Oy	933 AGAATATGGCCCAAGTTCTTGTTTATTATTCACAGACATTAAGTAGCATGCTGCCAGG 992
Db	2330 AGAAAAGAGACATTCCTAATTCACAGTCATTTTGGGAATTCCTGCTTAACCTCAAAAAT 2389
Oy	993 AGAAAAGAGACATTCCTAATTCACAGTCATTTTGGGAATTCCTGCTTAACCTCAAAAAT 1052
Db	2390 ATGGGAAAGACATGACGCTTCATGCCCCCTGACATCAAGAGATAGTGTAAAGAAC 2449
Oy	1053 ATGGGAAAGACATGACGCTTCATGCCCCCTGACATCAAGAGATAGTGTAAAGAAC 1112
Db	2450 AAGACATGTGTGTAT-AGAGACTCTCAATGATTTAGACAACTCAAAATACAGAGAA 2508
Oy	1113 AAGACATGTGTGTATAGAGACTCTCAATGATTTAGACAACTCAAAATACAGAGAA 1172
Db	2509 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATACATTTAAAGGGGAAAAA-- 2565
Oy	1173 AAGCAATGACTAGTAAATCATGTTAAAGGGGAAAAA-- 1232
Db	2566 CCCGACATCTCTCTCCGCCCTATTAATTGGAACAATAAGGTGAGGCTATCTC 2625
Oy	1233 CCCGACATCTCTCTCCGCCCTATTAATTGGAACAATAAGGTGAGGCTATCTC 1292
Db	2626 TACTTTCCTATCTGCGCAAGAAAGTGAAGAAATGGAGCTTTGGTTATTATGA 2685
Oy	1293 TACTTTCCTATCTGCGCAAGAAAGTGAAGAAATGGAGCTTTGGTTATTATGA 1352
Db	2686 TGGGACTGTAATTTGGTACAGATTTCTGAGGGCAATTTGGTAAATGATCAAAAGAC 2745
Oy	1353 TGGGACTGTAATTTGGTACAGATTTCTGAGGGCAATTTGGTAAATGATCAAAAGAC 1412
Db	2746 TTAATAATACGAGCGT 2761
Oy	1413 TTAATAATACGAGCGT 1428
RESULT	3
ID	PCT-US94-14920-1 STANDARD; DNA; UNC; 2763 BP.
AC	XXXXX
DT	01-JAN-1900
CC	Sequence 1, Application PC/TUS9414920.
CC	Sequence 1, Application PC/TUS9414920.
CC	GENERAL INFORMATION:
CC	APPLICANT: THE REGENTS OF THE UNIVERSITY

```
CC APPLICANT: OF CALIFORNIA
CC TITLE OF INVENTION: METHOD FOR DETECTION OF
CC TITLE OF INVENTION: METHYLTHIOADENOSINE PHOSPHATASE DEFICIENCY IN MANNA
ALIAN
CC TITLE OF INVENTION: CELLS
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Robbins, Berliner & Carson
CC STREET: 201 N. Figueroa Street, 5th floor
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90012
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/14920
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Berliner, Robert
CC REGISTRATION NUMBER: 20,121
CC REFERENCE/DOCKET NUMBER: 5555-287
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 213-977-1001
CC TELEFAX: 213-977-1003
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: methyladenosine phosphatase
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
S0 Sequence 2763 bp: 780 A; 505 C; 598 G; 803 T; 77 other;
Query Match 20.5%; Score 466; DB 10; Length 2763;
Best Local Similarity 99.0%; Pred. No. 0.00e+00;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Db 2270 AGAANAATGGCCACAGTTTCTGTTTATTATTAACAAGACATTAAAGTAGACATGCGTCCCAGG 2329
|||
QY 933 AGAATAATGGCCAGATTTCTGTGTTTATTATTAACAAGACATTAAAGTAGCAAGTGCGGCCAGG 992
2330 AGAAAAGAAGACATTCATTAATTCAGCATCTTTGGGAATTCCTGCTTAACCTGAATAAAT 2389
|||
QY 993 AGAAAAAGAAGACATTCATTAATTCAGCATCTTTGGGAATTCCTGCTTAACCTGAATAAAT 1052
2390 ATGGGAAAAGACATGCAGCTTCATGCCCCCTTGCTATCAAAGAGATATGTTGAAGAAAAGC 2449
QY 1053 ATGGGAAAAGACATGCAGCTTCATGCCCCCTTGCTATCAAAGAGATATGTTGAAGAAAAGC 1112
2450 AAGACATGTGTGAT-T-AGAGACTCCTCAATGATTTTAGCAACTTCAAAATACAGAAAGA 2508
|||
QY 1113 AAGACATGTGTGAT-T-AGAGACTCCTCAATGATTTTAGCAACTTCAAAATACAGAAAGA 1172
2509 AAGCAAAAGACTAGTAA-CATGTGGGAAAAAATATTACATTTTAAAGGGGAAAAA-- 2565
QY 1173 AAGCAAAAGACTAGTAAACATGTGGGAAAAAATATTACATTTTAAAGGGGAAAAA 1232
2566 CCCACCAATCTCTCTCCCTATTAATTTGCAACAATAAAGGTTGAGGGTATCTC 2625
QY 1233 CCCACCAATCTCTCTCCCTATTAATTTGCAACAATAAAGGTTGAGGGTATCTC 1292
Db .2626 TACTTCTATCTGCGCAAGAATGTGAGGAAGAATGGGACTCTTTGGTTATTATTGA 2685
|||||
```

```

0Y 1293 TACTTTCTATACGTGCCAAGATGTAGGAAGAAATGGACCTCTTGGTATTATTATGA 1352
Db 2686 TGGCAGCTGTAATATTGGTACAGTATTCTGTGAGGGCAATTGGTAAATGATCAAAAGAC 2745
OY 1353 TGGCGCTGTAATATTGGTACAGTATTCTGTGAGGGCAATTGGTAAATGATCAAAAGAC 1412
Db 2746 TTAATAATACGAGCT 2761
OY 1413 TTAATAATACGAGCT 1428

RESULT 4
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 5, Application US/08238163.
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POMELL, Ann
CC APPLICANT: STORZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL
CC DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourlie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc.feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name="deduced amino acid
CC OTHER INFORMATION: sequence of pcgip from bean."
CC
SQ Sequence 215 BP; 15 A; 8 C; 25 G; 26 T; 141 other;

Query Match 1.5%; Score 33; DB 5; Length 215;
Best Local Similarity 14.5%; Pred. No. 5.79e-06;
Matches 27; Conservative 75; Mismatches 82; Indels 2; Gaps 2;

Db 6 MSSSSVSRATASCNDKARKDNTTSMTTDCDCNRTWGCDDTDTYRVNDSGHNTYSSAN 65
CP 1081 AGCGCATGAAGCGTCGATGCTTCTCCCATATTTTTCGAA-GTAAAGACAGGAATCCCAA 1023
Db 66 YNYGNNVGAATHYTHYTHNVSGADSKVTVDSYNAGSTSSNGG-TDGNRSGADSYGSSK 124

```

Tue Feb 11 11:27:26 1997

US-08-674-311-1-01.rni

Page 4

[illegible]

RESULT 5 STANDARD; DNA; UNC; 215 BP.
ID US-08-238-163-5
AC xxxxxx
DT 01-JAN-1900
DE Sequence 5', Application US/08238163.
CC Sequence 5', Application US/08238163
CC Patent No. 556830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABARVITCH, John M.
CC APPLICANT: FORELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL
DISEASE

CC NUMBER OF SEQUENCES: 24
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Townsend and Townsend Kourie and Crew
 CC STREET: Steuart Street Tower, One Market Plaza
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: US
 CC ZIP: 94105-1493
 CC
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/258,163
 CC FILING DATE: 03-MAY-1994
 CC CLASSIFICATION: 800
 CC
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Bastian, Kevin D.
 CC REGISTRATION NUMBER: 34,774
 CC REFERENCE/DOCKET NUMBER: 2307E-540
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (415) 543-9600
 CC TELEFAX: (415) 543-5043
 CC INFORMATION FOR SEQ ID NO: 5:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 215 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: unknown
 CC MOLECULE TYPE: protein
 CC FEATURE:
 CC NAME/KEY: misc_feature
 CC LOCATION: 1..215
 CC OTHER INFORMATION: /standard_name="Deduced amino acid
 CC
 CC OTHER INFORMATION: sequence of PGIP from bean."
 CC
 CC Sequence 215 bp; 15 A; 8 C; 25 G; 26 T; 141 other;

	Query Match	Similarity	Score 31:	DB 5:	length 215:
	Best Local	11.98:	Pred. No. 7.15e-05:		
	Matches	20:	Conservative	75:	Indels 1: Gaps 1:
Db	6	MSSSVVSRTASCNDAKKGNTSSWTTDCCNRWGCYDTITRYNNDSDHNKYSAN	65		
		CGGGGCGAAGGATTATGTTCGGACACCTGGGGGCGGAGATGTTATACATACACACAGT	717		

[illegible]

```

ID      RESULT      6
AC      PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
DT      xxxxxx
DE      01-JAN-1900
DE      Sequence 100, Application PC/TUS9511934.
CC      Sequence 100, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC      APPLICANT: Cytogen Corporation
CC      TITLE OF INVENTION: Antigen Binding Peptides (AbIdies) From
CC      TITLE OF INVENTION: Peptide Libraries
CC      NUMBER OF SEQUENCES: 103
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Pennile & Edmonds
CC      STREET: 1155 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: USA
CC      ZIP: 10036
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patenrln release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US95/11934
CC      FILING DATE: 20-SEP-1995
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Mastrock, S. Leslie
CC      REGISTRATION NUMBER: 18,872
CC      REFERENCE/DOCKET NUMBER: 1101-196-228
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (212) 790-9090
CC      TELEFAX: (212) 869-9741/8864
CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 100:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 74 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (Genomic)
SQ      Sequence 74 BP: 6 A; 6 C; 1 G; 1 T; 60 other;

```

[illegible]

```

RESULT 7
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP
AC xxxxxx
DT 01-JUN-1900
DE Sequence 100, Application PC/TUS9511934
CC Sequence 100, Application PC/TUS9511934
GENERAL INFORMATION:

```

CC APPLICANT: Cyrogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtlides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEO ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 74 BP; 6 A; 6 C; 1 G; 1 T; 60 other:

Query Match 1.1%; Score 25; DB 11; Length 74;
Best Local Similarity 9.0%; Pred. No. 9.13e-02;
Matches 6; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

Db 6 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 65
Cc 469 CCGTCAAGACTGATCATATGACAAATGCGCGGTGATCTCCCTCAAGGA 410
Cc 66 ACCCAA 72
Cc 409 GCCACAA 403

RESULT 8
ID PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 99, Application PC/TUS9511934.
CC Sequence 99, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cyrogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtlides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEO ID NO: 99:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 75 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 75 BP; 1 A; 1 C; 7 G; 5 T; 61 other:

Query Match 1.1%; Score 25; DB 11; Length 75;
Best Local Similarity 12.5%; Pred. No. 9.13e-02;
Matches 9; Conservative 19; Mismatches 44; Indels 0; Gaps 0;

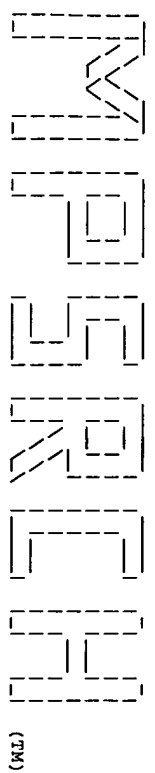
Db 4 AGNN 63
Cc 211 AGTTCTCTCTTAAATTTCTGATCATTGACGCGCTGTCCACCAATATTCACATCTT 152
Cc 64 NNBGGTTGGT 75
Cc 151 CACGCGGTGGT 140

RESULT 9
ID PCT-US95-11934-98 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 98, Application PC/TUS9511934.
CC Sequence 98, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cyrogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtlides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEO ID NO: 98:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid

CC LENGTH: 3214 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 220..2802
SQ Sequence 3214 BP: 897 A; 808 C; 792 G; 717 T; 0 other;
Query Match 1.1%; Score 24; DB 5; Length 3214;
Best Local Similarity 100.0%; Pred. No. 2.80e-01;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3144 TTGCTTTTAAAAA 3167
Oy 2237 TTGCTTTTAAAAA 2260
RESULT 15
ID PCT-US96-08014-269 STANDARD; DNA; UNC; 26 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 269, Application PC/TUS9608014.
CC Sequence 269, Application PC/TUS9608014
CC GENERAL INFORMATION:
CC APPLICANT: LARRY GOLD; NEBOLSA JANUIC; STEVEN RINGQUIST; NIKOS
CC APPLICANT: PARARATIS; PENLOPE J. TOOTHMAN
CC TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
CC TITLE OF INVENTION: LIGANDS TO TRANSFORMING GROWTH
CC TITLE OF INVENTION: FACTOR (TGF), PLATELET-DERIVED
CC TITLE OF INVENTION: GROWTH FACTOR (PDGF) AND HUMAN
CC TITLE OF INVENTION: KEROTINOCYTE GROWTH FACTOR (hKGF)
CC NUMBER OF SEQUENCES: 304
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Swanson & Bratschun, L.L.C.
CC STREET: 8400 E. Prentice Avenue, Suite 200
CC CITY: Englewood
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
CC COMPUTER: IBM compatible
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: WordPerfect 6.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/08014
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/458,423
CC FILING DATE: 02-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/458,424
CC FILING DATE: 02-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/455,594
CC FILING DATE: 03-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/455,591
CC FILING DATE: 03-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/479,725
CC FILING DATE: 07-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/479,783
CC FILING DATE: 07-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/618,693
CC FILING DATE: 20-MARCH-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Barry J. Swanson

CC REGISTRATION NUMBER: 33,215
CC REFERENCE/DOCKET NUMBER:
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (303) 793-3333
CC TELEFAX: (303) 793-3433
CC INFORMATION FOR SEQ ID NO: 269:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: RNA
CC FEATURE:
CC OTHER INFORMATION: All pyrimidines are 2'-fluoro (2'-F) modified
SQ Sequence 26 BP: 1 A; 0 C; 3 G; 0 T; 22 other;
Query Match 1.08; Score 22; DB 12; Length 26;
Best Local Similarity 16.7%; Pred. No. 2.45e-00;
Matches 4; Conservative 19; Mismatches 1; Indels 0; Gaps 0;
Db 1 RGRGRGAYRSBRSYXYYSYB 24
Oy 1320 AGGAAGAAATGGACTCTTGTT 1343

Search completed: Mon Feb 10 23:15:58 1997
Job time : 91 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Msrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Feb 10 23:16:19 1997; Maspar time 861.70 Seconds
Tabular output not generated. 953.554 Million cell updates/sec

Title: >US-08-674-311-1
Description: (1-2269) from US08674311.seq (1 of 2)
Perfect Score: 2269
N.A. Sequence: 1 GAATTCGCTCCGACACTGCT.....AAAAAAGCGAATTC 2269
Comp: CTTAAGCGGAGCGCTGACGA.....TTTTTTTTCGCTTAAAG

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 656168 segs, 181065524 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-pending
1:PCR91 2:PCR92 3:PCR93 4:PCR94 5:PCR95 6:PCR96 7:U75
8:U76 9:U77 10:U772 11:U781 12:U782 13:U791 14:U792
15:U800 16:U801 17:U802 18:U803 19:U804 20:U805 21:U806
22:U807 23:U808 24:U809 25:U810 26:U811 27:U812 28:U813
29:U814 30:U815 31:U816 32:U817 33:U818 34:U819 35:U820
36:U821 37:U822 38:U823 39:U824 40:U825 41:U826 42:U827
43:U828 44:U829 45:U830 46:U831 47:U832 48:U833 49:U834
50:U835 51:U836 52:U837 53:U838 54:U839 55:U840 56:U841
57:U842 58:U843 59:U844 60:U845 61:U846 62:U847 63:U848
64:U849 65:U85 66:U86 67:U87 68:U88 69:U89 70:U90
71:NEWU7 72:NEWU8

Statistics: Mean 10.868; Variance 5.284; scale 2.057
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description	Pred. No.
1	2269	100.0	US-08-674-	Sequence 1, Applicatio	0.00e+00
2	2269	100.0	US-60-000-	Sequence 1, Applicatio	0.00e+00
3	466	20.5	US-08-445-	Sequence 14, Applicati	0.00e+00
4	466	20.5	US-08-175-	Sequence 1, Applicatio	0.00e+00
5	466	20.5	US-08-443-	Sequence 14, Applicati	0.00e+00
6	466	20.5	US-08-227-	Sequence 14, Applicati	0.00e+00
7	466	20.5	US-08-227-	Sequence 14, Applicati	0.00e+00
8	466	20.5	US-08-220-	Sequence 14, Applicati	0.00e+00
9	185	8.2	US-08-196-	Sequence 5749, Applicat	6.39e-125
10	119	5.2	US-08-404-	Sequence 322, Applicat	4.84e-70

11	99	4.4	578	66	US-08-624-	Sequence 14, Applicati	7.12e-54
12	78	3.4	578	66	US-08-624-	Sequence 14, Applicati	2.59e-37
13	73	3.2	477	13	US-07-904-	Sequence 10, Applicati	1.86e-33
14	66	2.9	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
15	66	2.9	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
16	66	2.9	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
17	66	2.9	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
18	66	2.9	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
19	66	2.9	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
20	63	2.8	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
21	63	2.8	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
22	63	2.8	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
23	63	2.8	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
24	63	2.8	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
25	63	2.8	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
26	63	2.8	383	59	US-08-446-	Sequence 8, Applicatio	3.95e-28
27	64	2.8	7218	38	US-08-232-	Sequence 14, Applicati	1.26e-26
28	64	2.8	7218	38	US-08-232-	Sequence 14, Applicati	1.26e-26
29	61	2.7	339	66	US-08-617-	Sequence 8, Applicatio	2.18e-24
30	60	2.6	340	66	US-08-617-	Sequence 8, Applicatio	2.18e-24
31	58	2.6	477	13	US-07-904-	Sequence 9, Applicatio	1.20e-23
32	55	2.4	477	13	US-07-904-	Sequence 9, Applicatio	1.20e-23
33	50	2.2	340	66	US-08-617-	Sequence 9, Applicatio	5.59e-20
34	48	2.1	339	66	US-08-617-	Sequence 9, Applicatio	2.19e-16
35	41	1.8	7218	38	US-08-232-	Sequence 8, Applicatio	5.66e-15
36	41	1.8	7218	38	US-08-232-	Sequence 8, Applicatio	3.66e-10
37	36	1.6	234	13	US-07-904-	Sequence 14, Applicati	6.99e-07
38	37	1.6	1645	8	US-07-697-	Sequence 8, Applicatio	1.59e-07
39	37	1.6	1645	8	US-07-697-	Sequence 8, Applicatio	1.59e-07
40	33	1.5	234	13	US-07-904-	Sequence 8, Applicatio	5.40e-05
41	34	1.5	235	13	US-07-904-	Sequence 8, Applicatio	1.29e-05
42	34	1.5	965	53	US-08-388-	Sequence 22, Applicati	5.40e-05
43	33	1.4	222	32	US-07-904-	Sequence 58, Applicati	2.22e-04
44	32	1.4	235	13	US-07-904-	Sequence 6, Applicatio	2.22e-04
45	32	1.4	235	13	US-07-904-	Sequence 6, Applicatio	2.22e-04

ALIGNMENTS

RESULT 1
ID US-08-674-311-1 STANDARD; DNA; INC; 2269 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Applicatio US/08674311
CC Sequence 1, Applicatio US/08674311
CC GENERAL INFORMATION: US/08674311
CC APPLICANT: Olopad; Olufunmilayo I.
CC TITLE OF INVENTION: METHYLTHIOADENOSINE PHOSPHORYLASE
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE IN
CC TITLE OF INVENTION: EARLY DIAGNOSIS AND TREATMENT OF
CC TITLE OF INVENTION: PROLIFERATIVE DISORDERS
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P.O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: United States of America
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/674,311
CC FILING DATE: Concurrently Herewith
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 60/000,831
CC FILING DATE: 02-JUL-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:

Tue Feb 11 11:27:27 1997

US-08-674-311-1-01.rmp

Page 2

CC NAME: Kitchell, Barbara S.
CC REGISTRATION NUMBER: 33,928
CC REFERENCE/DOCKET NUMBER: ARSB:509
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (512) 474-7577
CC TELEFAX: (512) 418-3000
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2269 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 122..970
CC Sequence 2269 bp: 725 A; 407 C; 490 G; 647 T; 0 other;

Query Match 100.0%; Score 2269; DB 66; Length 2269;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GAATTCGCTCCGCACTGCTCCTCCGCGCAGTGAAGTTGGACACACCGCTCTGTG 60
1 GAATTCGCTCCGCACTGCTCCTCCGCGCAGTGAAGTTGGACACACCGCTCTGTG 60
Qy 61 GCGCGTTGGTCCCTTAAGTCCGAGCGCTGCCACTGCAGATTCTTCCGTCGAGA 120
61 GCGCGTTGGTCCCTTAAGTCCGAGCGCTGCCACTGCAGATTCTTCCGTCGAGA 120
Db 121 CATGGCTCTGGACACACACCGCGGTGAAGTTGGAAATTTGGAGACAGGCTT 180
121 CATGGCTCTGGACACACACCGCGGTGAAGTTGGAAATTTGGAGACAGGCTT 180
Qy 121 CATGGCTCTGGACACACACCGCGGTGAAGTTGGAAATTTGGAGACAGGCTT 180
121 CATGGCTCTGGACACACACCGCGGTGAAGTTGGAAATTTGGAGACAGGCTT 180
Db 181 GGATGATCCAGAAATTTGAAGAGAACTGAATAATATGTGATCTCCATTGGCAA 240
181 GGATGATCCAGAAATTTGAAGAGAACTGAATAATATGTGATCTCCATTGGCAA 240
Qy 181 GGATGATCCAGAAATTTGAAGAGAACTGAATAATATGTGATCTCCATTGGCAA 240
181 GGATGATCCAGAAATTTGAAGAGAACTGAATAATATGTGATCTCCATTGGCAA 240
Db 241 GCATCTGATGCTTAAATTTGGGGAAGTAAAAATGTGATTCATCTCTTGGCAAG 300
241 GCATCTGATGCTTAAATTTGGGGAAGTAAAAATGTGATTCATCTCTTGGCAAG 300
Qy 241 GCATCTGATGCTTAAATTTGGGGAAGTAAAAATGTGATTCATCTCTTGGCAAG 300
241 GCATCTGATGCTTAAATTTGGGGAAGTAAAAATGTGATTCATCTCTTGGCAAG 300
Db 301 GCATGGAAGGACACACCATCATGCTTCAAAAGTCAACTACAGAGGAAATCTGGGC 360
301 GCATGGAAGGACACACCATCATGCTTCAAAAGTCAACTACAGAGGAAATCTGGGC 360
Qy 301 GCATGGAAGGACACACCATCATGCTTCAAAAGTCAACTACAGAGGAAATCTGGGC 360
301 GCATGGAAGGACACACCATCATGCTTCAAAAGTCAACTACAGAGGAAATCTGGGC 360
Db 361 TTGGAAGGAAGAGGCTGTACACATGTCTATGTACACACAGCTGTGCTTGAAGGA 420
361 TTGGAAGGAAGAGGCTGTACACATGTCTATGTACACACAGCTGTGCTTGAAGGA 420
Qy 361 TTGGAAGGAAGAGGCTGTACACATGTCTATGTACACACAGCTGTGCTTGAAGGA 420
361 TTGGAAGGAAGAGGCTGTACACATGTCTATGTACACACAGCTGTGCTTGAAGGA 420
Db 421 GGAGATTCAACCCGGGAGATTTGTCTATTATGATCAGTTCAATTGACAGACCTATGAG 480
421 GGAGATTCAACCCGGGAGATTTGTCTATTATGATCAGTTCAATTGACAGACCTATGAG 480
Qy 421 GGAGATTCAACCCGGGAGATTTGTCTATTATGATCAGTTCAATTGACAGACCTATGAG 480
421 GGAGATTCAACCCGGGAGATTTGTCTATTATGATCAGTTCAATTGACAGACCTATGAG 480
Db 481 ACCCTAGTCTCTATGATGAAGATCTTGTGCCAGAGAGAGTGTCCATATTCCAAT 540
481 ACCCTAGTCTCTATGATGAAGATCTTGTGCCAGAGAGAGTGTCCATATTCCAAT 540
Qy 481 ACCCTAGTCTCTATGATGAAGATCTTGTGCCAGAGAGAGTGTCCATATTCCAAT 540
481 ACCCTAGTCTCTATGATGAAGATCTTGTGCCAGAGAGAGTGTCCATATTCCAAT 540
Db 541 GCGGACCCGTTTGGCCCAAAACGAGAGAGTCTTATGAGAGTGTGAAGAGTGAAG 600
541 GCGGACCCGTTTGGCCCAAAACGAGAGAGTCTTATGAGAGTGTGAAGAGTGAAG 600
Qy 541 GCGGACCCGTTTGGCCCAAAACGAGAGAGTCTTATGAGAGTGTGAAGAGTGAAG 600
541 GCGGACCCGTTTGGCCCAAAACGAGAGAGTCTTATGAGAGTGTGAAGAGTGAAG 600
Db 601 ACTCCGCTGCTCAAAAGGAGCAATGTCTCAATCGAGGAGCTCTTTTACCTCCG 660
601 ACTCCGCTGCTCAAAAGGAGCAATGTCTCAATCGAGGAGCTCTTTTACCTCCG 660
Qy 601 ACTCCGCTGCTCAAAAGGAGCAATGTCTCAATCGAGGAGCTCTTTTACCTCCG 660
601 ACTCCGCTGCTCAAAAGGAGCAATGTCTCAATCGAGGAGCTCTTTTACCTCCG 660
Db 661 GCGGAAAGCTTCAATGTTCCGACCTGGGGGGGATGTATCAATGACCAAGTTCC 720
661 GCGGAAAGCTTCAATGTTCCGACCTGGGGGGGATGTATCAATGACCAAGTTCC 720
Qy 661 GCGGAAAGCTTCAATGTTCCGACCTGGGGGGGATGTATCAATGACCAAGTTCC 720
661 GCGGAAAGCTTCAATGTTCCGACCTGGGGGGGATGTATCAATGACCAAGTTCC 720
Db 721 AGAGAGGTTCTTGCTAAGAGAGCTGAATTTGTTACGCAATATCGCATGGGAGACGA 780
721 AGAGAGGTTCTTGCTAAGAGAGCTGAATTTGTTACGCAATATCGCATGGGAGACGA 780
Qy 721 AGAGAGGTTCTTGCTAAGAGAGCTGAATTTGTTACGCAATATCGCATGGGAGACGA 780
721 AGAGAGGTTCTTGCTAAGAGAGCTGAATTTGTTACGCAATATCGCATGGGAGACGA 780

Db 781 TTATGACTGCTGGAAGAGACGAGAGAGAGTTCGGTGGACCGGCTTAAAGACCT 840
781 TTATGACTGCTGGAAGAGACGAGAGAGAGTTCGGTGGACCGGCTTAAAGACCT 840
Qy 781 TTATGACTGCTGGAAGAGACGAGAGAGAGTTCGGTGGACCGGCTTAAAGACCT 840
781 TTATGACTGCTGGAAGAGACGAGAGAGAGTTCGGTGGACCGGCTTAAAGACCT 840
Db 841 GAAAGAAAACGCTAATTAAGCCAAAGCTTACCTCAGTACATCCAGTAGAGCTC 900
841 GAAAGAAAACGCTAATTAAGCCAAAGCTTACCTCAGTACATCCAGTAGAGCTC 900
Qy 841 GAAAGAAAACGCTAATTAAGCCAAAGCTTACCTCAGTACATCCAGTAGAGCTC 900
841 GAAAGAAAACGCTAATTAAGCCAAAGCTTACCTCAGTACATCCAGTAGAGCTC 900
Db 901 CACAGAAAGTGTGAGAAACCTCCATTAACCTGAAGATATGAGCCAGTTTCTTTAT 960
901 CACAGAAAGTGTGAGAAACCTCCATTAACCTGAAGATATGAGCCAGTTTCTTTAT 960
Qy 901 CACAGAAAGTGTGAGAAACCTCCATTAACCTGAAGATATGAGCCAGTTTCTTTAT 960
901 CACAGAAAGTGTGAGAAACCTCCATTAACCTGAAGATATGAGCCAGTTTCTTTAT 960
Db 961 ACCAGACATTAAGTAGATGCTGCCAGAGAGAGAGACATTTCAATCCAGTCA 1020
961 ACCAGACATTAAGTAGATGCTGCCAGAGAGAGAGACATTTCAATCCAGTCA 1020
Qy 961 ACCAGACATTAAGTAGATGCTGCCAGAGAGAGAGACATTTCAATCCAGTCA 1020
961 ACCAGACATTAAGTAGATGCTGCCAGAGAGAGAGACATTTCAATCCAGTCA 1020
Db 1021 TTTTGGGAATTCCTGCTTAACCTGAAAAAATATGGGAAAGACATGCTTCATGCC 1080
1021 TTTTGGGAATTCCTGCTTAACCTGAAAAAATATGGGAAAGACATGCTTCATGCC 1080
Qy 1021 TTTTGGGAATTCCTGCTTAACCTGAAAAAATATGGGAAAGACATGCTTCATGCC 1080
1021 TTTTGGGAATTCCTGCTTAACCTGAAAAAATATGGGAAAGACATGCTTCATGCC 1080
Db 1081 TTGCTATCAAGAGATATGTTGAAGAAAGCAAGACATTTGTGTATTAGACCTCTG 1140
1081 TTGCTATCAAGAGATATGTTGAAGAAAGCAAGACATTTGTGTATTAGACCTCTG 1140
Qy 1081 TTGCTATCAAGAGATATGTTGAAGAAAGCAAGACATTTGTGTATTAGACCTCTG 1140
1081 TTGCTATCAAGAGATATGTTGAAGAAAGCAAGACATTTGTGTATTAGACCTCTG 1140
Db 1141 AATGATTAGACAACTTCAAAATACAGAGAAAGCAATGACTAGTAAACATGGGAA 1200
1141 AATGATTAGACAACTTCAAAATACAGAGAAAGCAATGACTAGTAAACATGGGAA 1200
Qy 1141 AATGATTAGACAACTTCAAAATACAGAGAAAGCAATGACTAGTAAACATGGGAA 1200
1141 AATGATTAGACAACTTCAAAATACAGAGAAAGCAATGACTAGTAAACATGGGAA 1200
Db 1201 AAAATATTAATTTAAGGGGAAAAAAACCCACATTCCTCTCCCTATTA 1260
1201 AAAATATTAATTTAAGGGGAAAAAAACCCACATTCCTCTCCCTATTA 1260
Qy 1201 AAAATATTAATTTAAGGGGAAAAAAACCCACATTCCTCTCCCTATTA 1260
1201 AAAATATTAATTTAAGGGGAAAAAAACCCACATTCCTCTCCCTATTA 1260
Db 1261 AATTGCAACAATAAGGGGAGGTAATCTTCTTCTATCTGCAAGAGATGTA 1320
1261 AATTGCAACAATAAGGGGAGGTAATCTTCTTCTATCTGCAAGAGATGTA 1320
Qy 1261 AATTGCAACAATAAGGGGAGGTAATCTTCTTCTATCTGCAAGAGATGTA 1320
1261 AATTGCAACAATAAGGGGAGGTAATCTTCTTCTATCTGCAAGAGATGTA 1320
Db 1321 GGAAGAAATGGACTTTGTTATTTATGATGCGACTGTAATTTGTAAGATTTCT 1380
1321 GGAAGAAATGGACTTTGTTATTTATGATGCGACTGTAATTTGTAAGATTTCT 1380
Qy 1321 GGAAGAAATGGACTTTGTTATTTATGATGCGACTGTAATTTGTAAGATTTCT 1380
1321 GGAAGAAATGGACTTTGTTATTTATGATGCGACTGTAATTTGTAAGATTTCT 1380
Db 1381 GGAAGGCAATTTGTTAAATGATCAAAAGACTTAAATAGGAGCTCTTGGTCTG 1440
1381 GGAAGGCAATTTGTTAAATGATCAAAAGACTTAAATAGGAGCTCTTGGTCTG 1440
Qy 1381 GGAAGGCAATTTGTTAAATGATCAAAAGACTTAAATAGGAGCTCTTGGTCTG 1440
1381 GGAAGGCAATTTGTTAAATGATCAAAAGACTTAAATAGGAGCTCTTGGTCTG 1440
Db 1441 GGAAGCTATCATCTAGCAATTTCTTTAAACATATCAAGATGATCAAAAGATTA 1500
1441 GGAAGCTATCATCTAGCAATTTCTTTAAACATATCAAGATGATCAAAAGATTA 1500
Qy 1441 GGAAGCTATCATCTAGCAATTTCTTTAAACATATCAAGATGATCAAAAGATTA 1500
1441 GGAAGCTATCATCTAGCAATTTCTTTAAACATATCAAGATGATCAAAAGATTA 1500
Db 1501 TATTAAGAAAGGCTTTAATTAATGATTAATTAATTAATTAATTAATTAATTA 1560
1501 TATTAAGAAAGGCTTTAATTAATGATTAATTAATTAATTAATTAATTAATTA 1560
Qy 1501 TATTAAGAAAGGCTTTAATTAATGATTAATTAATTAATTAATTAATTAATTA 1560
1501 TATTAAGAAAGGCTTTAATTAATGATTAATTAATTAATTAATTAATTAATTA 1560
Db 1561 ATCCCTGCAATTTGAGATTAATGATGCTTAAAGACAGGTTGCTATATATGAA 1620
1561 ATCCCTGCAATTTGAGATTAATGATGCTTAAAGACAGGTTGCTATATATGAA 1620
Qy 1561 ATCCCTGCAATTTGAGATTAATGATGCTTAAAGACAGGTTGCTATATATGAA 1620
1561 ATCCCTGCAATTTGAGATTAATGATGCTTAAAGACAGGTTGCTATATATGAA 1620
Db 1621 CTGAATAATCTTTTGCATATTTCAATGTCTTAAAGACAGGTTGCTATATGAA 1680
1621 CTGAATAATCTTTTGCATATTTCAATGTCTTAAAGACAGGTTGCTATATGAA 1680
Qy 1621 CTGAATAATCTTTTGCATATTTCAATGTCTTAAAGACAGGTTGCTATATGAA 1680
1621 CTGAATAATCTTTTGCATATTTCAATGTCTTAAAGACAGGTTGCTATATGAA 1680
Db 1661 GTGAAGAAAGATATGATGATTAATGATGATGATGATGATGATGATGATGATG 1740
1661 GTGAAGAAAGATATGATGATTAATGATGATGATGATGATGATGATGATGATG 1740
Qy 1661 GTGAAGAAAGATATGATGATTAATGATGATGATGATGATGATGATGATGATG 1740
1661 GTGAAGAAAGATATGATGATTAATGATGATGATGATGATGATGATGATGATG 1740
Db 1741 TATACAAAAAACTAGAAAGAAATATATTAACCTTTGATTTTGGGGAGGATA 1800
1741 TATACAAAAAACTAGAAAGAAATATATTAACCTTTGATTTTGGGGAGGATA 1800
Qy 1741 TATACAAAAAACTAGAAAGAAATATATTAACCTTTGATTTTGGGGAGGATA 1800
1741 TATACAAAAAACTAGAAAGAAATATATTAACCTTTGATTTTGGGGAGGATA 1800
Db 1801 CTGGGATTAATTTTATTTCTTGAATCTTCTGCTTCACATTTTCTACAGTAAAT 1860
1801 CTGGGATTAATTTTATTTCTTGAATCTTCTGCTTCACATTTTCTACAGTAAAT 1860
Qy 1801 CTGGGATTAATTTTATTTCTTGAATCTTCTGCTTCACATTTTCTACAGTAAAT 1860
1801 CTGGGATTAATTTTATTTCTTGAATCTTCTGCTTCACATTTTCTACAGTAAAT 1860
Db 1861 ATATCAAAATATGTAAGGCGCTTAAATAAAGTGAATTTGAAGAAAGATCCAGTTCTGA 1920
1861 ATATCAAAATATGTAAGGCGCTTAAATAAAGTGAATTTGAAGAAAGATCCAGTTCTGA 1920

QY	1861	ATAATCAAAATAGTAAGGAGCGCCGTAATAATTAAGTGGATTGTGAAGAATCAGTCTTGA	1920
Db	1921	AAACACGTTCTTCGTAAATGAACACAAATTAGTGGTAATATTAAGTGAATCTATT	1980
QY	1921	AAACACGTTCTTCGTAAATGAACACAAATTAAGTGGTAATTAAGTGAATCTATT	1980
Db	1981	TAAAGGAGTTACATCTTATTCTGCTAAAGAGAGCATATTGATTCCTGTACACTCAGA	2040
QY	1981	TAAAGGAGTTACATCTTATTCTGCTAAAGAGAGCATATTGATTCCTGTACACTCAGA	2040
Db	2041	ACAACTACTTGGGCTGCAACAAGCTTCTGTAGAAAGAGCTAGCGTATATAGTTAACTGA	2100
QY	2041	ACAACTACTTGGGCTGCAACAAGCTTCTGTAGAAAGAGCTAGCGTATATAGTTAACTGA	2100
Db	2101	AAGTTAACTATTAAAGACATAATGACATTTATGGTATCTGATATTTAAAAAGTA	2160
QY	2101	AAGTTAACTATTAAAGACATAATGACATTTATGGTATCTGATATTTAAAAAGTA	2160
Db	2161	ATGTGAGCTTCCTCTTTTATAGAGTTAAATTATTTATACGAGTGGTAATTTGTGCTT	2220
QY	2161	ATGTGAGCTTCCTCTTTTATAGAGTTAAATTATTTATACGAGTGGTAATTTGTGCTT	2220
Db	2221	TTAATAAGTGAAGCCTTGGTTTAAAAAAGGGAATTC	2269
QY	2221	TTAATAAGTGAAGCCTTGGTTTAAAAAAGGGAATTC	2269

RESULT 2
 US-60-000-831-1 STANDARD; DNA; UNC; 2269 BP.
 ID
 DT xxxxxx
 AC 01-JAN-1990
 DE Sequence 1, Application US/60000831.
 CC Sequence 1, Application US/60000831
 CC GENERAL INFORMATION:
 CC APPLICANT: Olopade, Olufunmilayo I.
 CC APPLICANT: Bohlander, Stefan K.
 CC APPLICANT: Dreyling, Martin H.
 CC APPLICANT: Le Beau, Michelle M.
 CC APPLICANT: Stadler, Walter M.
 CC APPLICANT: Gursky, Susan
 CC APPLICANT: Espinosa, Iri, Rafael
 CC APPLICANT: Sween, Lise W.
 CC APPLICANT: Hagos, Fltsum
 CC APPLICANT: Pomykala, Helen M.
 CC APPLICANT: Adeyanju, M.O.
 CC TITLE OF INVENTION: METHYLTIOADENOSINE PHOSPHORYLASE
 CC TITLE OF INVENTION: METHODS AND COMPOSITIONS
 CC NUMBER OF SEQUENCES: 2
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Arnold, White & Durkee
 CC STREET: P.O. Box 4433
 CC CITY: Houston
 CC STATE: Texas
 CC COUNTRY: United States of America
 CC ZIP: 77210
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/60/000,831
 CC FILING DATE: Concurrently Herewith
 CC CLASSIFICATION:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Kitchell, Barbara S.
 CC REGISTRATION NUMBER: 33,928
 CC REFERENCE/DOCKET NUMBER: ARCD:156/KIT
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (512) 418-3000
 CC TELEFAX: (512) 474-7577
 CC INFORMATION FOR SEQ ID NO: 1:
 CC

```

CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 2269 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: other nucleic acid
CC      DESCRIPTION: /desc = "DNA"
CC      FEATURE:
CC      NAME/KEY: CDS
CC      LOCATION: 122..970
CC      Sequence 2269 BP: 725 A; 407 C; 490 G; 647 T; 0 other;
SQ
Query Match      100.0%; Score 2269; DB 68; Length 2269;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2269; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

841 CAGAGATGTCGAAACCCCTACCTGAGAAATGAGGCTC 900
 Db 901 CAGAGATGTCGAAACCCCTACCTGAGAAATGAGGCTC 960
 Qy 901 CAGAGATGTCGAAACCCCTACCTGAGAAATGAGGCTC 960
 Db 961 ACCAGACATTTAAAGTACGATGCTCCAGAGAAAGAGATTCATTCAGTCA 1020
 Qy 961 ACCAGACATTTAAAGTACGATGCTCCAGAGAAAGAGATTCATTCAGTCA 1020
 Db 1021 TTTTGGAAATTCCTGCTTAACCTGAAAAAATATGAGAAAGATGCACTTCATCC 1080
 Qy 1021 TTTTGGAAATTCCTGCTTAACCTGAAAAAATATGAGAAAGATGCACTTCATCC 1080
 Db 1081 TTGCTATCAAGAGATGTTGTAAGAAAGACAGCATTTGTTATTAAGAGACTCTG 1140
 Qy 1081 TTGCTATCAAGAGATGTTGTAAGAAAGACAGCATTTGTTATTAAGAGACTCTG 1140
 Db 1141 AATGATTTAGCAACTCAAAATACAGAAAGCAATGACTAGTAATGAGGAA 1200
 Qy 1141 AATGATTTAGCAACTCAAAATACAGAAAGCAATGACTAGTAATGAGGAA 1200
 Db 1201 AAAATATTTACATTTTAAAGGGGAAAAAAACCCGACATCTCTGCCCCATTTAA 1260
 Qy 1201 AAAATATTTACATTTTAAAGGGGAAAAAAACCCGACATCTCTGCCCCATTTAA 1260
 Db 1261 ATTGCAACAATAAAGGAGTGAAGGATATCTCTTCTTACTGCCAAGATGTA 1320
 Qy 1261 ATTGCAACAATAAAGGAGTGAAGGATATCTCTTCTTACTGCCAAGATGTA 1320
 Db 1321 GGAACAAATGAGACTCTTGTGTTATTAATGATGACGACTGTAATTTGATGATTTCT 1380
 Qy 1321 GGAACAAATGAGACTCTTGTGTTATTAATGATGATGATGATGATGATTTCT 1380
 Db 1381 GAGAGGCAATTTGTAATAATGATCAAAAGCACTTAAATAGGAGCTCTTGTGCTG 1440
 Qy 1381 GAGAGGCAATTTGTAATAATGATCAAAAGCACTTAAATAGGAGCTCTTGTGCTG 1440
 Db 1441 GGAAGCTTACATTCAGCAATTTCTTTAAACATTCAGAGATGCAATCAAGATTA 1500
 Qy 1441 GGAAGCTTACATTCAGCAATTTCTTTAAACATTCAGAGATGCAATCAAGATTA 1500
 Db 1501 TATATAAAGAGGAGTGTAAATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Qy 1501 TATATAAAGAGGAGTGTAAATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Db 1561 ATCCCTGCAATTTGAGGTAATATGCTAGTATTAATCTAGATTTGCAATCCCA 1620
 Qy 1561 ATCCCTGCAATTTGAGGTAATATGCTAGTATTAATCTAGATTTGCAATCCCA 1620
 Db 1621 CTGAAAAATCCTTTGCAATTTCAATGCTTAAAGACAGCGTTCCTATATATGAA 1680
 Qy 1621 CTGAAAAATCCTTTGCAATTTCAATGCTTAAAGACAGCGTTCCTATATATGAA 1680
 Db 1681 GTGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Qy 1681 GTGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Db 1741 TATACAAAAAACTAGAAAGAAATATATATATATATATATATATATATATATATAT 1800
 Qy 1741 TATACAAAAAACTAGAAAGAAATATATATATATATATATATATATATATATATAT 1800
 Db 1801 CTGGGATATTTTATTTTCTTTGATCTCTCTGCTTACATTTTCTTACAGTAA 1860
 Qy 1801 CTGGGATATTTTATTTTCTTTGATCTCTCTGCTTACATTTTCTTACAGTAA 1860
 Db 1861 ATATCAATATGTAAGGCGCTTAAAGAAAGATGTAAGAAAGATGTAAGAAAGATGTA 1920
 Qy 1861 ATATCAATATGTAAGGCGCTTAAAGAAAGATGTAAGAAAGATGTAAGAAAGATGTA 1920
 Db 1921 AAACACTGTTCTGTAATGAGCAGATTTAGTTGTAATTTAAAGGATGTCATT 1980
 Qy 1921 AAACACTGTTCTGTAATGAGCAGATTTAGTTGTAATTTAAAGGATGTCATT 1980

Qy 1921 AAACACTGTTCTGTAATGAGCAGATTTAGTTGTAATTTAAAGGATGTCATT 1980
 Db 1981 TAAGGAGATTCATCTTTTCTGTAAGAGAGATCAATGATTTCTGTAAGTCA 2040
 Qy 1981 TAAGGAGATTCATCTTTTCTGTAAGAGAGATCAATGATTTCTGTAAGTCA 2040
 Db 2041 ACAGTACTGCTGCTTCTTCTGTAAGAGAGATCAATGATTTCTGTAAGTCA 2100
 Qy 2041 ACAGTACTGCTGCTTCTTCTGTAAGAGAGATCAATGATTTCTGTAAGTCA 2100
 Db 2101 AAGTTTAACTATTTAAAGACTTAAAGCAATTTATGATGATGATTTTAAAGTA 2160
 Qy 2101 AAGTTTAACTATTTAAAGACTTAAAGCAATTTATGATGATGATTTTAAAGTA 2160
 Db 2161 ATGAGACTCTCTCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 Qy 2161 ATGAGACTCTCTCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 Db 2221 TTAATTAAGTGAAGCTGCTTTTAAAGAAAGAAAGAAAGAAAGAAAGTA 2269
 Qy 2221 TTAATTAAGTGAAGCTGCTTTTAAAGAAAGAAAGAAAGAAAGAAAGTA 2269

RESULT 3
 ID US-08-445-648-14 STANDARD; DNA; UNC. 2763 BP.
 AC xxxxxx
 DT 01-JAN-1900
 DE Sequence 14, Application US/08445648.
 CC Sequence 14, Application US/08445648
 CC GENERAL INFORMATION:
 CC APPLICANT: CARSON, DENNIS A.
 CC APPLICANT: NOBORI, TSUTOMU
 CC TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR
 CC TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESSI
 ON AND CANCER
 CC TITLE OF INVENTION: TREATMENT
 CC NUMBER OF SEQUENCES: 18
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Fish & Richardson P.C.
 CC STREET: 4225 Executive Square, Suite 1400
 CC CITY: LA Jolla
 CC STATE: California
 CC COUNTRY: US
 CC ZIP: 92037
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/445,648
 CC FILING DATE: 22-MAY-1995
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/227,800
 CC FILING DATE: 14-APR-1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: HOWELL, STACY L.
 CC REGISTRATION NUMBER: 34,842
 CC REFERENCE/DOCKET NUMBER: 07340/023001
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 619/678-5070
 CC TELEFAX: 619/678-5098
 CC INFORMATION FOR SEQ ID NO: 14:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 2763 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: DNA (genomic)
 CC IMMEDIATE SOURCE:
 CC CLONE: Methylothiodenosine Phosphorylase (genomic)
 CC FEATURE:

CC	NAME/KEY:	CDS	LOCATION:	1..2763
SC	Sequence	2763 BP:	780 A; 505 C; 598 G; 803 T; 77 other:	
CC	Query Match	20.5%;	Score 466; DB 59; Length 2763;	
CC	Best Local Similarity	99.0%;	Pred. No. 0.00e+00;	
SC	Matches	491; Conservative	0; Mismatches 1; Indels 4; Gaps 3;	
Db	2270	AGAAATATGCCCCAGTTTCTGTGTTTATTATACCAAGACATTTAAAGTAGCATGCGTCCGACAG	2329	
QY	933	AGAAATATGCCCCAGTTTCTGTGTTTATTATACCAAGACATTTAAAGTAGCATGCGTCCGACAG	992	
Db	2330	AGAAAAGAAAGACATCTCAATTCGATCATCTTTGGGAATTCCTGCTTAACCTGAAAAAAT	2389	
QY	993	AGAAAAGAAAGACATCTCAATTCGATCATCTTTGGGAATTCCTGCTTAACCTGAAAAAAT	1052	
Db	2390	ATGGGAAAGACATGCGAGCTTTCATCCCTGCGCTATCAAGAGATATGTTGTAAGAAAGAC	2449	
QY	1053	ATGGGAAAGACATGCGAGCTTTCATCCCTGCGCTATCAAGAGATATGTTGTAAGAAAGAC	1112	
Db	2450	AAGACATTGTGTGAT-AGAGACTCCCTCAATGATTAGACAACCTCAAAATACAGAGAA	2508	
QY	1113	AAGACATTGTGTGATAGAGACTCCCTCAATGATTAGACAACCTCAAAATACAGAGAA	1172	
Db	2509	AAGCAATGACTAGTAA-CATGTGGGAAAAAATATTACCTTTTAAAGGGGAAAAA--	2565	
QY	1173	AAGCAATGACTAGTAAATCATGTGGGAAAAAATATTACCTTTTAAAGGGGAAAAA--	1232	
Db	2566	CCCCACATTCCTCTCTCCCTATTAATTTGCAACAATAAAGGTGGAGGCTAATCTC	2625	
QY	1233	CCCCACATTCCTCTCTCTCCCTATTAATTTGCAACAATAAAGGTGGAGGCTAATCTC	1292	
Db	2626	TACTTTCCTATACGCCAAGAAATGTGAGGAAGAAATGGAGCTCTTGTGTTATTATTGA	2685	
QY	1293	TACTTTCCTATACGCCAAGAAATGTGAGGAAGAAATGGAGCTCTTGTGTTATTATTGA	1352	
Db	2686	TGCACCTGTAATTTGTCACAGTATTTCTGAGGCGCAATTTGTTAAATGCATCAAAAGAC	2745	
QY	1353	TGCACCTGTAATTTGTCACAGTATTTCTGAGGCGCAATTTGTTAAATGCATCAAAAGAC	1412	
Db	2746	TTAAATAATACGAGCT	2761	
QY	1413	TTAAATAATACGAGCT	1428	
RESULT	4	US-08-176-855-1 STANDARD; DNA; UNC; 2763 BP.		
ID	US-08-176-855-1	STANDARD; DNA; UNC; 2763 BP.		
AC	xxxxxx			
DT	01-JAN-1900			
CC	Sequence 1, Application	US/08176855.		
CC	Sequence 1, Application	US/08176855		
CC	GENERAL INFORMATION:			
CC	APPLICANT:	Nobori, Tsutomu		
CC	APPLICANT:	Carson, Dennis A.		
CC	APPLICANT:	Takadayashl, Kenji		
CC	TITLE OF INVENTION:	METHOD FOR DETECTION OF		
CC	TITLE OF INVENTION:	METHYLLHIOADENOSINE PHOSPHATASE DEFICIENCY IN MAMM		
CC	ALLIAN			
CC	TITLE OF INVENTION:	CELLS		
CC	NUMBER OF SEQUENCES:	5		
CC	CORRESPONDENCE ADDRESS:			
CC	ADDRESSEE:	Spensley Horn Judas & Lubitz		
CC	STREET:	1880 Century Park East, Suite 500		
CC	CITY:	Los Angeles		
CC	STATE:	California		
CC	COUNTRY:	USA		
CC	ZIP:	90067		
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE:	Floppy disk		
CC	COMPUTER:	IBM PC compatible		
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS		
CC	SOFTWARE:	PatentIn Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:			

CC	APPLICATION NUMBER:	US/08/176,855
CC	FILING DATE:	
CC	CLASSIFICATION:	435
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	Howells, Stacy L.
CC	REGISTRATION NUMBER:	34,842
CC	REFERENCE/DOCKET NUMBER:	EP3057
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	(619) 455-5100
CC	TELEFAX:	(619) 455-5110
CC	INFORMATION FOR SEQ ID NO:	1:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	2763 base pairs
CC	TYPE:	nucleic acid
CC	STRANDEDNESS:	single
CC	TOPOLOGY:	linear
CC	MOLECULE TYPE:	DNA (genomic)
CC	IMMEDIATE SOURCE:	
CC	CLONE:	methyladenosine phosphatase
CC	FEATURE:	
CC	NAME/KEY:	CDS
CC	LOCATION:	1..2763
CC	Sequence	2763 bp: 780 A; 505 C; 598 G; 803 T; 77 other;
CC	Query Match	20.5%; Score 466; DB 32; Length 2763;
CC	Best Local Similarity	99.0%; Pred. No. 0.00e+00;
CC	Matches	491; Conservative 0; Mismatches 1; Indels 4; Gaps 3
Db	2270	AGAAATATGGCCAGTTTCTGTTTATTATTCACAAACATTAAGTAGCATGGCTCCAGG 2329
Qy	933	AGAAATATGGCCAGTTTCTGTTTATTATTCACAAACATTAAGTAGCATGGCTCCAGG 992
Db	2330	AGAAAGAGAGACATCTAATTCACAGCTATTTGGGAATTCCTGCTTAACCTGAAGAAAAAT 2389
Qy	993	AGAAAGAGAGACATCTAATTCACAGCTATTTGGGAATTCCTGCTTAACCTGAAGAAAAAT 1052
Db	2390	ATGGGAAAGACATGACGCTTCATGCCCCCTGCTCATCAAGAGATATGTTGAAGAAAGAC 2449
Qy	1053	ATGGGAAAGACATGACGCTTCATGCCCCCTGCTCATCAAGAGATATGTTGAAGAAAGAC 1112
Db	2450	AAGACATTTGTGTTT-AGAGATCTCTCAATGATTTAGACACTTCAAAAATACGAAGAA 2508
Qy	1113	AAGACATTTGTGTTTAGAGACCTCCGAAAGATTAGACAACTTCAAAAATACGAAGAA 1172
Db	2509	AAGCAATGACTAGTAA-CATGTGGGAAAAATATTACATTTTAAGGGGGGAAAAAAA-- 2565
Qy	1173	AAGCAATGACTAGTAAATGAGAAAAATTTTACATTTTAAGGGGGGAAAAAAA 1232
Db	2566	CCCCACCATTTCTTCTCCCCCTATTAAATTTGGACAAATAAAGGTGAGGTTATCTC 2625
Qy	1233	CCCCACCATTTCTTCTCCCCCTATTAAATTTGGACAAATAAAGGTGAGGTTATCTC 1292
Db	2626	TACTTTCCTATTAATCTCCAAAGATGTGAGGAAGAAATGGGACTTTTGGTTATTATTGA 2685
Qy	1293	TACTTTCCTATTAATCTCCAAAGATGTGAGGAAGAAATGGGACTTTTGGTTATTATTGA 1352
Db	2686	TGCGACTGTAATGTGTCAGATTTTCGAGGGCAATTGGTAAATGCATCAAAAGAC 2745
Qy	1353	TGCGACTGTAATGTGTCAGATTTTCGAGGGCAATTGGTAAATGCATCAAAAGAC 1412
Db	2746	TTAAAAATACGAGCT 2761
Qy	1413	TTAAAAATACGAGCT 1428
RESULT	5	
ID	US-08-443-880-14	STANDARD; DNA; UNC; 2763 BP.
AC	xxxxxx	
DT	01-JAN-1900	
DE	Sequence 14, Application US/08443880.	
CC	Sequence 14, Application US/08443880.	
CC	GENERAL INFORMATION:	
CC	APPLICANT:	CARSON, DENNIS A.

Db	2330	AGAAAAGACATCTTAATTCAGACATTTGGGAATTCGCGTTAACTGGAAAAAT	2389
Qy	993	AGAAAAGACATTTCTTAATTCAGATCATTTTGGATTTCTGCTTAAGTAAAAAT	10522
Db	2390	ATGGAAAGACATGCACGCTTTCATGCGCTTGCCCTATCAAGAGTATGTTGAAGAAC	2449
Qy	1053	ATGGAAAGACATGCACGCTTTCATGCGCTTGCCCTATCAAGAGTATGTTGAAGAAC	11122
Db	2450	AAGCATTTGTGTAT-AGAGACTCCTCAATGATTTAGACAACCTCAAAATACAGAGA	25080
Qy	1113	AAGCATTTGTGTATGATGAGACTCCTGATGATTTAGACAACCTCAAAATACAGAGA	11722
Db	2509	AAGCAATGCTAGTA-CATGAGGAAAAAATATACATTTTAAGGGGAAAAAA--	25555
Qy	1173	AAGCAATGCTAGTATAACATGTGGAAAAAATATACATTTTAAGGGGAAAAAA	12322
Db	2566	CCCCACATTCCTCTTCCCTCATTTAAATTTGCAACAATAAGGTTGAGGATATCTC	26255
Qy	1233	CCCCACATTCCTCTTCCCTCATTTAAATTTGCAACAATAAGGTTGAGGATATCTC	12322
Db	2626	TACTTTCCTACTGCGCAAGAAATGTAGAGAGAAATGGGACTCTTTGGTATTTATGA	26855
Qy	1293	TACTTTCCTACTGCGCAAGAAATGTAGAGAGAAATGGGACTCTTTGGTATTTATGA	13522
Db	2686	TGCACGTAAATGTGATGACGATTTTCTGAGGGCAATTTGGTAAATGCACAAAAGC	27455
Qy	1353	TGCACGTAAATGTGATGACGATTTTCTGAGGGCAATTTGGTAAATGCACAAAAGC	14122
Db	2746	TTAAAAATACGAGCT 2761	
Qy	1413	TTAAAAATACGAGCT 1428	
RESULT	7		
ID	US-08-227-800-14	STANDARD; DNA; UNC; 2763 BP.	
AC	xxxxxx		
DT	01-JAN-1900		
DE	Sequence 14, Application US/08227800.		
CC	Sequence 14, Application US/08227800		
CC	GENERAL INFORMATION:		
CC	APPLICANT: CARSON, DENNIS A.		
CC	APPLICANT: NOBORI, TSUTOMU		
CC	TITLE OF INVENTION: METHODS FOR TREATMENT AND DETECTION OF		
CC	TITLE OF INVENTION: SUSCEPTIBILITY TO, AND PROGRESSION OF, CANCER		
CC	NUMBER OF SEQUENCES: 14		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: SPENSLEY HORN JUBAS & LOBITZ		
CC	STREET: 1880 Century Park East, Suite 500		
CC	CITY: Los Angeles		
CC	STATE: California		
CC	COUNTRY: US		
CC	ZIP: 90067		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/227,800		
CC	FILING DATE:		
CC	CLASSIFICATION: 435		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: HOWELLS, STACY L.		
CC	REGISTRATION NUMBER: 34,842		
CC	REFERENCE/DOCKET NUMBER: PD3459		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 619/455-5100		
CC	TELEFAX: 619/455-5110		
CC	INFORMATION FOR SEQ ID NO: 14:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 2763 base pairs		
CC	TYPE: nucleic acid		
CC	STRANDEDNESS: single		

CC	TOOLOGY: linear
CC	MOLECULE TYPE: DNA (genomic)
CC	IMMEDIATE SOURCE:
CC	CLONE: Methylthiodenosine phosphorylase (genomic)
CC	FEATURE:
CC	NAME/KEY: CDS
CC	LOCATION: 1..2763
CC	Sequence 2763 BP: 780 A; 505 C; 598 G; 803 T; 77 other:
S0	
	Query Match 20.5%; Score 466; DB 37; Length 2763;
	Best Local Similarity 99.0%; Pred. No. 0.00e+00;
	Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
D0	AGATATGCCCCAGTTTCTGTTTATTATACCAGACATTAAGTAGCATGGCTGCCAGG 2329
0Y	AGAAATATGCCCCAGTTTCTGTTTATTATACCAGACATTAAGTAGCATGGCTGCCAGG 992
D0	AGAAAAGAGACATTCTTAATCCAGCATTTTGGGAATTCCTGTTAAGTGAAGAAAAT 2389
0Y	AGAAAAGAGACATTCTTAATCCAGCATTTTGGGAATTCCTGTTAAGTGAAGAAAAT 1052
D0	ATGGGAAAGACATGACGTTTCATGACCCCTTGCCTATCAAGAGATGTTGTAAGAAAG 2449
0Y	ATGGGAAAGACATGACGTTTCATGACCCCTTGCCTATCAAGAGATGTTGTAAGAAAG 1112
D0	2450 AAGACATTGTGTAT-AGAGACTCTCAATGATTTAGACACTTCAAAATACAGAGAA 2508
0Y	1113 AAGACATTGTGTATAGAGACTCTCAATGATTTAGACACTTCAAAATACAGAGAA 1172
D0	2509 AAGCAATGACTACTGA-CATGTGGGAAAAAATTTTCATTTTAAGGGGAAAAAATA 2565
0Y	1173 AAGCAATGACTACTGAATCATGTGGGAAAAAATTTTCATTTTAAGGGGAAAAAATA 1232
D0	2566 CCCACCATCTCTCTCCGCCCTATTAAATTTGCAACAATAAAGGGTGAAGGTAATCTC 2625
0Y	1233 CCCACCATCTCTCTCCGCCCTATTAAATTTGCAACAATAAAGGGTGAAGGTAATCTC 1292
D0	2626 TACTTCTCTATAGCCAAAGAAATGTAGGAAGAAATGGGACTTTGGTTATTATTTGA 2685
0Y	1293 TACTTCTCTATAGCCAAAGAAATGTAGGAAGAAATGGGACTTTGGTTATTATTTGA 1352
D0	2686 TGGCAGCTAAATTTGTCACGATATTTCTGAGAGCAATTTGTTAAATGATCAAAAAGAC 2745
0Y	1353 TGGCAGCTAAATTTGTCACGATATTTCTGAGAGCAATTTGTTAAATGATCAAAAAGAC 1412
D0	2746 TTAATAATACGAGCT 2761
0Y	1413 TTAATAATACGAGCT 1428
RESULT	8
ID	US-08-220-691-5097 STANDARD; DNA; UNC; 495 BP.
AC	xxxxxx
DT	01-JAN-1900
DE	Sequence 5097, Application US/08220691.
CC	Sequence 5097, Application US/08220691.
CC	GENERAL INFORMATION:
CC	APPLICANT: Haseltine, William, et. al.
CC	TITLE OF INVENTION: Human Genes, Sequences, and
CC	TITLE OF INVENTION: ExpressionProducts
CC	NUMBER OF SEQUENCES: 12049
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
CC	ADDRESSEE: Cecchi, Stewart, & Olstein
CC	STREET: 6 Becker Farm Road
CC	CITY: Roseland
CC	STATE: New Jersey
CC	COUNTRY: USA
CC	ZIP: 07068
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
CC	COMPUTER: HP Vectra 486/33
CC	OPERATING SYSTEM: MSDOS version 5.0

Tue Feb 11 11:27:27 1997

US-08-674-311-1-01.rnp

Page 8

```
CC SOFTWARE: ASCII text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/220.691
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Olstein, Elliot
CC REGISTRATION NUMBER: 24,025
CC REFERENCE/DOCKET NUMBER: 325800-71
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (201) 994-1700
CC TELEFAX: (201) 994-1744
CC INFORMATION FOR SEQ ID NO: 5097:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 495 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC Sequence 495 bp; 97 A; 96 C; 127 G; 131 T; 44 other.
```

Query Match	11.48;	Score 258;	DB 37;	Length 495;
Best Local Similarity	84.88;	Pred. No. 4.72e-187;		
Matches	352;	Conservative	0;	Mismatches 52; Indels 11; Gaps 11;

Db	14	GNCAGTATGNAAGTTTGGACACACGACACCCCTCTGTGGGGCGCTTGGTTCCTTAGTCCAG	73
Qy	28	GCCCACTG - AGSTTGGACACACACCGCTCTGGGCGCTTGGTTCCTTAGTCCAG	86
Db	74	CGCTGCCACACTGACGATTCCTTCCGTGCCGAGACATAGGCTCTGGACACACACACG	133
Qy	87	CGCTGCCACACTGACGATTCCTTCCGTGCCGAGACATAGGCTCTGGACACACACACG	148
Db	134	CGCTGTAAATTTGNNAAATATTTGGGGGAACGGCGTGAATGATCCAGAA - TTTTAAAGG	192
Qy	147	CGCTG - AAATTTG - AATPATTTGGGGAACGGCGTGAATGATCCAGAAATTTTAAAGG	204
Db	193	AAGAACTGCCAAATATATGAGTACTCCATTTGGCAAGCATCTGAATGCTTAATTTGGG	253
Qy	205	AAGAACTGCCAAATATATGAGTACTCCATTTGGCAAGCATCTGAATGCTTAATTTGGG	264
Db	253	GGAAGTAAAAAATGTGTGATTTGCGTCTCTTGGCAAGGCGATGTTAGNAGNACCAT	312
Qy	265	G - AAGATAAAAAATGTGTGATTTGCGTCTCTTGG - CAA - GGCAATGGAAGGACACACACCAT	321
Db	313	CANGGCCCTTAAAGGTCATTAATCCAGGGGANNTTTGGGTTTNNAAAGAA - AGGGTTN	371
Qy	322	CATG - CCTTCAAAGG - CAACATCCAGGGCAACATCTGGGCTTTGAAGGAAGAGGGCGCT	379
Db	372	NAACCTTTTNGNAGCCNAAANTTNNNGTNTTGGGGGGGGTNTTNAACC	426
Qy	380	ACCAATGATATAGTGACACAGCTT - GTGGCTCCTTTGAGGAGGAATTCACACC	433

RESULT	-	9	
ID	US-08-196-363-5749	STANDARD	DNA; UNC; 205 BP
AC	xxxxxx		
DT	01-JAN-1900		
DE	Sequence 5749	Application	US/08196363.
CC	Sequence 5749	Application	US/08196363
CC	GENERAL INFORMATION:		
CC	APPLICANT:	Kerlavage, Anthony	
CC	APPLICANT:	Fraser, Claire M.	
CC	APPLICANT:	Kirkness, Eben	
CC	APPLICANT:	Sutton, Granger	
CC	APPLICANT:	Cocayne, Jeannine	
CC	APPLICANT:	Lin, Li-Ing	
CC	APPLICANT:	Fitzgerald, Lisa	
CC	APPLICANT:	Adams, Mark	
CC	APPLICANT:	Lee, Normal	
CC	APPLICANT:	Fuldner, Rebecca	

```

CC APPLICANT: Fleischmann, Robert
CC APPLICANT: Bult, Carol
CC APPLICANT: Blake, Judy
CC APPLICANT: White, Owen
CC APPLICANT: Clayton, Rebecca
CC APPLICANT: Pellegrino, Susan
CC TITLE OF INVENTION: Human Genes, Sequences, and Expression
CC TITLE OF INVENTION: Products
CC NUMBER OF SEQUENCES: 16818
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart, &
CC ADDRESS: Olstein
CC STREET: 6 Becker Farm Road
CC CITY: Roseland
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07068
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 Inch, 1.4mb storage
CC COMPUTER: HP Vectra 486/33
CC OPERATING SYSTEM: MSDOS version 5.0
CC SOFTWARE: ASCII Text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/196,363
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Olstein, Elliot
CC REGISTRATION NUMBER: 12,025
CC REFERENCE/DOCKET NUMBER: 325800-91
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (201) 994-1700
CC TELEFAX: (201) 994-1744
CC INFORMATION FOR SEQ. ID NO.: 5749:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 205 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
Qq Sequence 205 bp; 72 A; 26 C; 31 G; 70 T; 6 other;

```

Query Match	8.2%	Score 185;	DB 34;	Length 205;
Best Local Similarity	94.7%	Pred. No. 6.39e-125;		
Matches	195;	Conservative	0;	Mismatches 10; Indels 1; Gaps 1

Db	1	ATTAAACANCTGATTCCTTCGATCTGAGGTAAATATGTTTGTATTAAT-TAGA	59
Db	1546	ATTGAACAACACTGAATCCCTTCGATCTGAGGTAAATATGCTTGTATATATCTACA	1605
Qy			
Db	60	TGTGTAAACANCAACGAAATATCTTTNNCATATTTTCATGTCCTAAACACACGGT	119
Qy	1606	TGTGTAAATACGCAACCTGAAATATCTTTTTCGATATTTTCATATGCTCTAAACACACGGT	1665
Db	120	TGCNCTATATATGAGGTGAAAAAAGCATATGTTAGCATTTATATAGTACTAGTTTCTCTT	179
Db	1666	TGCTCTATATATGAAGTGAATAAAAGGATATGGTACGATTTTATATAGTACTAGTTTCTCTT	1725
Qy			
Db	180	AAATATCTATGTAATATATACAAAAA	205
Db	1726	AAATATCTATGTAATATATACAAAAA	1751
Qy			

RESULT - 10
ID US-08-404-891-322 STANDARD: DNA: UNC: 137 BP
AC xxxxxx
DT 01 JAN 1900
DE Sequence 322, Application US/08404891.
CC Sequence 322, Application US/08404891.
CC GENERAL INFORMATION:
CC APPLICANT: Stuart, Susan G.

Tue Feb 11 11:27:27 1997

US-08-674-311-1-01.rnp

Page 10

Dh 539 RKSRSNTTTRVKSAAVHSGKRDYRD 563
QY 886 ACCTCAGATAGGCTCCACAGATG 910

RESULT 12
ID US-08-624-190-14 STANDARD; DNA; UNC; 578 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08624190.
CC Sequence 14, Application US/08624190.
CC GENERAL INFORMATION:
CC APPLICANT: KAVANAGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC APPLICANT: WILLIAMS PH.D., LEWIS T.
CC TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
CC TITLE OF INVENTION: 5-PHOSPHATASE
CC NUMBER OF SEQUENCES: 38
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHIRON CORPORATION
CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/624,190
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.003
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ Sequence 578 BP; 37 A; 11 C; 62 G; 65 T; 403 other;

Query Match 3.4%; Score 78; DB 66; Length 578;
Best Local Similarity 13.0%; Pred. No. 2,59e-37;
Matches 36; Conservative 193; Mismatches 177; Indels 3; Gaps 3;

Dh 31 DKVKAHGRHSRYTVKASGKMKVDYSGKRSKSGSKSHKSKKNVYTKRKYVADSK 90
CP 1080 GGGCATGAAAGCTGATCTTCCCATATTTTTCAGTTAAGTAAGCAGATCCCAAAA 1021

Dh 91 RGCMMKNKSDMTGTWMMGNNAKTSMSKSGKTRDSDADYHYGYGTDGSKMSTSVTVAH 150
CP 1020 TGACGTGAATAGATGCTCTTTCTCTCTGCGCAGCCATGCTACTTATGCTGTGT 961

Dh 131 TWNRVVAHNSHCTDNVKTGANTGNKGVNSGTSGVNSHTSGSKRRNNYMRACD 210
CP 960 AATAAAGAGAAAGTGGCCATATTTCTTCAAGTTAAGGAGGCTTCTGACCATCTGTG 901

Dh 211 KKSNTHTWMDNTRVDWTATKRYADSHDRKRVHTATYRTRD-KYATTKATGMYNS 269
CP 900 GACCTATCTGAGTAGTAGTGCAGCAGTAAGCTTTGCTTATTAAGGTTTCTTTC 841

Dh 270 WCDRWKYSYVWVSYSSTSDMTSDSVATAGYTSVSKNSGTVDSRCATATKSKYNS 329
CP 840 AGGGCTTTAAGACCCGGTCCACGGAACATGCTTCTCTGCTCTTCCAGCAGTCAATA 781

Dh 330 SCVSKSGSGGVYKTRKSDYHSHKSSDSYGCARATYTTTHGTGHTSGTKRKYD 389
CP 780 TCTGTCCCATGGCGATAGTGGTAAACAAATTCAGCCTCTTACGAAGAACCACTCT 721

Dh 390 VKTR-DSSGKTSTSHDMWVTSRACSGSTNNYGVGMVHKSDT-AWSYDKDSCRG 447
CP 720 GGACTGTGTCATGTTGATACATCCGCCCCCGAGGTGGAGACATGACCTTCTGCC 661

Dh 448 STGSKKSTANR 458
CP 660 CGGAGCTAA 650

RESULT 13
ID US-07-904-074A-10 STANDARD; DNA; UNC; 477 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 10, Application US/07904074A.
CC Sequence 10, Application US/07904074A.
CC GENERAL INFORMATION:
CC APPLICANT: John E. Shively
CC APPLICANT: Anna Wu
CC APPLICANT: Rainer Fischer
CC APPLICANT: Roy Paxton
CC APPLICANT: Y. H. Joy Yang
CC TITLE OF INVENTION: Chimeric Anti-CEA Antibody
CC NUMBER OF SEQUENCES: 10
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: City of Hope
CC STREET: 1500 East Duarte Road
CC CITY: Duarte
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 91010-0269
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3M Double Density 5 1/4" diskette
CC COMPUTER: Wang PC
CC OPERATING SYSTEM: MS-DOS (R) Version 3.30
CC SOFTWARE: Microsoft (R)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/904,074A
CC FILING DATE: 19920615
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA: None
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Irons, Edward S.
CC REGISTRATION NUMBER: 16,541
CC REFERENCE/DOCKET NUMBER: None
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202) 785-6938
CC TELEFAX: (202) 785-5351
CC INFORMATION FOR SEQ ID NO: 10:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 477
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: Single Stranded
CC TOPOLOGY: Unknown
CC MOLECULE TYPE: Nucleic Acid
CC ANTI-SENSE: Not Applicable
CC FRAGMENT TYPE: Not Applicable
CC ORIGINAL SOURCE: Synthetically Prepared
CC IMMEDIATE SOURCE: Synthetically Prepared
CC POSITION IN GENOME: None
CC FEATURE: None
CC PUBLICATION INFORMATION: None
SQ Sequence 477 BP; 23 A; 12 C; 35 G; 35 T; 372 other;

Query Match 3.2%; Score 73; DB 13; Length 477;
Best Local Similarity 8.5%; Pred. No. 1.86e-33;
Matches 39; Conservative 154; Mismatches 263; Indels 1; Gaps 1;

```
Db 5 HRRNTNNNGNNNNNNVNVKGVNCKNVNNGGNGVKNKNGSNKNSCAAGNTNNSYAMSW 64
Oy 21 CACTCCCGGCGAGTGAAGTTGGCAGCCAGCCCTCTGTGGCTGTGGTTCCCTTAGT 80
Db 65 VRNTNNKNNVNVNNSSDGNTNVDSYKGNTRYSRDNRNNNNNNNSNRSNTAYYCA 124
Oy 81 CCGGAGCGGCGCGCCACTGAGATTCCTTCCCGCTGACAGATGGCCTCGGACACCA 140
Db 125 RNDYGGGGNGVNGTNGTNAVSASTRKSNVNNNANSKSTSGGTAAAGCNGVNDYNNNY 184
Oy 141 CCACCCCGGAGATGATGAATGATGATGAGGAGGCTGATGATCCAGAAATTTAG 200
Db 185 TVSNNGANTSGVHTNNVNVNNSGNSNSSVYVNVSSNTNTYNNVNNKNSNTYDKR 244
Oy 201 AAGGAAGAACTGAAATAATATGATGATCTGCAATTTGGCAAGCATCGATGATTTT 260
Db 245 VNKRSCDHTHCNCCNANNNGNNGSVNNNNNNKHKDRTMNNRTNNVTCVVDASHNDNNVK 304
Oy 261 TGGGGAAGATAAAAAATGTTGATTCATCTCTCTGCAAGCATGGAAGCAGACACCA 320
Db 305 NNNYVDGVNVNNAKTRNNNNNNSTYRVSVNTVNHNDMNNNGKNYCKYNNKANNNNK 364
Oy 321 TCATGCTTCAAGGTCAAC-TACACAGCGACATCTGGGCTTGAAGGAAGAGCGGT 379
Db 365 TNSKAGNNNNNNVYNNNSRDNNTRNNVNTCNVKNYSNDNAVMNNSNGNNNNKTT 424
Oy 380 ACACATGATCTAGTACACAGCTGTGGCTCTGTGAGGAGAGATTCAGCCCGGAT 439
Db 425 NNVNDSGNSNNYKNTVDRKRNNGNNGSVYHNA 461
Oy 440 ATGTCTATATTGATCAGTTCATTCAGACAGACCACTA 476

RESULT 14
ID US-08-446-939-8 STANDARD; DNA; UNC; 383 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 8, Application US/08446939.
CC Sequence 8, Application US/08446939
CC GENERAL INFORMATION:
CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF- Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millita Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,939
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
```

```
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI91-09FX
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 383 BP; 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match 2.9%; Score 66; DB 59; Length 383;
Best Local Similarity 11.9%; Pred. No. 3,95e-28;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;

Db 1 MTSSTGCHSGAAGHVTWRASTHVKSNNNDVTDNNGAVKCKDYRSTCDNKSCHSNC 60
Oy 697 ATCCGCCCCCGAGTCCGGAACATGAAGCTTCTGCCCGGAGCTAAACGAGGTCCCTC 638
Db 61 TSCVGVAVWRKNDNTVCHDKYHDDAASKMKKKGTWCSCSDCNDNSYNTSDVYTG 120
Oy 637 GATTGTGACATTTGCTCCCTTTGAGTGCCAGGAGCTCTAGCTTTTGTAGCATCTC 579
Db 121 SGVAVYCYRVNKRSTWTGTRKMSHCADRSSTCANNNHNTDTVGRAVYRAKN 180
Oy 578 T-AAGAACCTCTCTCTTTGGGCAAAACGGCTGACCATGGAATA-TGGACACTCC 521
Db 181 TSTVAVKTHYASWKRKSDNKHNTARKTKYTWAAAGNTRHVSWMRNVSSARGSH 240
Oy 520 TCTGGCACAAGATGACTTTCATCATGAGAGACTGAGTCTCATGTGCTCTGCAAT 461
Db 241 HSDHTGRKVRHDKRSNPNKNDCCDGSRGYSVDANSVGTATAYMVSNNNAKTD 300
Oy 460 GAACATCATCATATATACATATATCGCCGGCTGAATCTCTCTCT-CAAGAGCCCAAG 402
Db 301 VYSMAVMTSRCAVGVKDYGSVRDVSMMKDNVRDGRTRNSWMHGNVCTTCMDHART 360
Oy 401 CTGTGTCTACTATGACATGTGTACAGCCCTCTCTCTCAAGCCGAGATGTTGGCTGT 342
Db 361 A 361
Oy 341 A 341

RESULT 15
ID US-08-446-936-8 STANDARD; DNA; UNC; 383 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 8, Application US/08446936.
CC Sequence 8, Application US/08446936
CC GENERAL INFORMATION:
CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF- Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millita Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
```

Tue Feb 11 11:27:27 1997

US-08-674-311-1-01.rnp

Page 12

CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,936
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WH191-09FY
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-9240
CC TELEFAX: 617-861-9340
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC Sequence 383 BP: 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match 2.9% Score 66; DB 59; length 383;
Best Local Similarity 11.9% Pred. No. 3,958-28;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;
Db 1 MTSSGVCHGSAAGAHVHTASTHVSVDNYTDNNGAVKCKCDVYSTCDNKSCHMSNCS 60
Cp 697 ATCCGCCGCCAGGTGGGACATGAGCTTCTGCCCGGAGCTAAACAGAGTCCCTC 638
Db 61 TSCVYCAVWRKNDNTVCHDKYHDAASKCMKKKGTMCSCSDCNDNSYNTSNDVYTG 120
Cp 637 GATGTGACCATGTGCTCCCTTGAGTGCGACCGAGCTCCTAGCTCTTAGCAGCTC-TA 579
Db 121 SGVAVYCYVNRKSTWTGTRKMSHCHADRSSTCANNNHNTDYGKRAVYKARKN 180
Cp 578 T-AAGAACCTCTCGCTTTGGGGCAAAAGGCTCAGCCATGGAATA-TGGCACACTCC 521
Db 181 TSTVAVYVDHYASWKDSDNKNHTARKTGKVTAAKNGYTRHVSMDRNVGSSARBSH 240
Cp 520 TCTGGCACAGATGCTTCATCATGAGAGCTAGGCTCATATGAGTCTCTCAAT 461
Db 241 HSDHTCGRKVNRKSNVNDTCCDGSRGYSVDANSVGTARYMAVSRMNNASTD 300
Cp 460 GAAGTGCATCAATATGATGCGGGGCTGAATCTCCTCCT-CAAGGAGCCACAAG 402
Db 301 VYSMAVMTSRCAVGVKDYKSVRDVNSKDNVRDRTNRSSVNHGVCYTCWHDART 360
Cp 401 CTGTGTCTACTATGACATGTAGACCCCTCTCTCAAGCCAGATGTTCCGCTGT 342
Db 361 A 361
Cp 341 A 341

Search completed: Mon Feb 10 23:43:38 1997
Job time: 1639 secs.